PTO 2001-3941

WIPO, International Patent Application
WO98/49284

METHODS FOR DETECTING AND ISOLATING NUCLEAR TRANSPORT PROTEINS

[Kakuikō Tanpakushitsu no Kenshutsu Oyobi Tanri Hōhō]

Nobuhide UEKI et al.

UNITED STATES PATENT AND TRADEMARK OFFICE
Washington, D.C.		September 2001

Translated by: Schreiber Translations, Inc.

Country	:	World Intellectual Property Organization
Document No.	:	WO98/49284
Document Type	:	International Application Published Pursuant to Patent Cooperation Treaty
Language	:	Japanese
Inventor	:	Nobuhide UEKI, Teruhisa NOGUCHI
Applicant	:	Helix Research Institute
IPC[6]	:	C12N 15/09, C12Q 1/68
Filing Date	:	April 27, 1998
Publication Date	:	November 5, 1998
Priority Information	:	Tokugan Hei 9/124795 April 28, 1997 Tokugan Hei 9/309686 October 24, 1997
Foreign Language Title	:	Kakuikō Tanpakushitsu no Kenshutsu Oyobi Tanri Hōhō
English Title	:	METHODS FOR DETECTING AND ISOLATING NUCLEAR TRANSPORT PROTEINS

1/[1]

Specification

Methods For Detecting and Isolating Nuclear Transport Proteins

Technical Field

The present invention relates to methods of detecting and isolating nuclear transport proteins, and falls into the field of genetic engineering, more particularly gene cloning.

Prior Art

Various transcription factors, nuclear receptors, signal transfer factors, chromatin receptors, and the like are known as nuclear transport proteins. These proteins interact directly or indirectly with specific DNA regions in the vicinity of the end of intracellular signal transfer cascades to control gene expression, replication of DNA, and the like, and as a result, determine the behavior of cells. Accordingly, the isolation of the genes of these nuclear transport proteins and analysis of their functions is thought to be highly important from the viewpoints of elucidating various vital phenomena and developing new drugs.

However, no specific method of comprehensively cloning cDNA coding for nuclear transport proteins has been developed; the general methods that have been applied in cloning techniques thus far are used. That is, when there is some information relating to a protein the cloning of which is being attempted -- for example, when there is a sequence that is stored at the amino acid level (Lichtsteiner, S., Proc. Natl. Acad. Sci., 1993, 90: 9673-9677), an interacting DNA sequence is already known (Sanz, L., Mol. Cell. Biol., 1995, 15: 3164-3170; made by Clontech Co., Matchmaker One-Hybrid System), or an interacting protein is already known -- a cDNA library is cloned based on that information in these methods. However, in such cases, screening is possible only within an extremely limited range.

/2

It is known, for example, that the "Two-Hybrid System" (Gyuris, J., Cell, 1993, 75: 791-803; Golemis, E.A., Current Protocols in Molecular Biology (John Wiley & Sons, Inc.), 1996, Ch. 20.0 and 20.1) developed in recent years as a method of isolating interacting proteins can be employed by using as bait a protein already known to be present in the nucleus and thereby indirectly screen for cDNA coding proteins that interact with that protein (Jordan, K.L., Biochemistry, 1996, 35: 12,320-12,328). However, it cannot be directly employed as a method of screening cDNA coding for proteins that have transport activity into the nucleus. Further, even when employing bait in the form

[1] Numbers in the margin indicate pagination in the original.

of a protein known to be present in the nucleus, since it is not known whether transport into the nucleus occurs through interaction in the cytoplasm or through actual interaction in the nucleus, there is also the possibility that cDNA coding for proteins other than nucleoproteins will also end up being isolated. Thus, an arduous confirmation operation is necessary to determine whether or not the isolated cDNA codes for a nuclear transport protein. Further, since the "Two-Hybrid System" indicates interaction between proteins, there is also a problem in that the proteins obtained by screening end up being limited to just proteins capable of interacting with the protein employed as bait.

When it is impossible to obtain information relating to the targeted protein in the manner described above, it is necessary to extract nuclear fractions from the cell, refine the targeted protein therefrom by a method employing functions such as specific biological activity possessed by that protein as indicators, and screen a cDNA library based on sequence information relating to the protein obtained (Ostrowski, J., J. Biol. Chem., 1994, 269: 17,626-17,634). However, some nuclear transport proteins have extremely low expression levels, often necessitating the expenditure of considerable time and effort to refine, with some of them being nearly impossible to refine.

/3

Disclosure of the Invention

The present invention has as its object to provide methods of readily and efficiently detecting and isolating DNA coding for peptides having nuclear transport capability.

One example of a nuclear transport protein is transcription factor. The transcription factor of eukaryotic organisms has the functions of migrating into the nucleus and inducing the expression of a specific gene by interacting with the promoter region of that specific gene. The nuclear migration ability of transcription factor is attributed to a nuclear migration signal present in the transcription factor. The present inventors focused on the two characteristics of transcription factor having the ability to migrate into the nucleus and the ability to activate transcription in a specific gene and conducted extensive research into resolving this issue. As a result, they discovered that when the region having the nuclear transportability in transcription factor was eliminated, an unknown peptide was introduced in place thereof, and the protein thus obtained was expressed within the cell, if the unknown peptide in the fused protein had the ability to migrate into the nucleus, it was transported with the fused protein into the nucleus, acted on a

particular promoter region, and was thought to induce the expression of a specific gene downstream therefrom. Further, if the unknown peptide in the fused protein did not have the ability to migrate into the nucleus, the fused protein did not migrate into the nucleus and was thought not to induce the expression of a specific gene in a downstream region of the promoter. That is, by means of a protein in which an unknown peptide had been fused into transcription factor not having the ability to migrate into the nucleus, it was thought to be possible to determine whether or not the unknown peptide in the fused protein had the ability to migrate into a nucleus based on indication in the form of inducement of the expression of a particular gene downstream from the promoter.

Accordingly, based on this idea, the present inventors actually prepared fused DNA of test DNA and DNA coding for transcription factor from which the region having the ability to migrate into the nucleus had been removed, introduced into a eukaryotic host maintaining in the nucleus a promoter region activated by the binding of transcription factor and a reporter gene the expression of which was induced by activation of that promoter region, and detected the expression of the reporter gene. As a result, they discovered that when DNA coding for a peptide having the ability to migrate into the nucleus was employed as test DNA, expression of the reporter gene was induced, and when DNA coding for a peptide not having the ability to migrate into the nucleus was employed as test DNA, expression of the report gene was not induced.

/4

The present inventors further prepared a library of cDNA coding for fused proteins of transcription factor from which the region having the ability to migrate into nucleuses had been removed and other peptides, and introduced these into cells to screen cDNA coding for peptides having nuclear transportability employing the expression of the reporter as indicator. As a result, the present inventors discovered that most of the known cDNA isolated from the cDNA library coded for proteins thought to have nuclear transportability.

That is, the present invention relates to methods of readily and efficiently detecting and isolating DNA coding for peptides having nuclear transportability using the properties of transcription factor, and more particularly, relates to:

(1) A method of detecting the nuclear transportability of a peptide coded for by test DNA, characterized in that fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is introduced into a eukaryotic

host having in its nucleus a promoter region that is activated by binding of the transcription factor and a reporter gene connected downstream from the promoter region, and detecting expression of the reporter gene;

(2) The method described in (1) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA bonding domain, and a transcription activation domain;

(3) The method described in (1) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and a GAL4 transcription activation domain, and wherein the promoter region that is activated by binding of the transcription factor is the promoter region of a GAL1 gene in which the operator sequence has been replaced with the LexA operator sequence;

(4) The method described in (3) wherein the nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5;

(5) The method described in any of (1)-(4) wherein the reporter gene is the LEU2 and/or β-galactosidase gene;

(6) A method of isolating DNA coding for a peptide having nuclear transportability characterized in that fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is introduced into a eukaryotic host having in its nucleus a promoter region that is activated by the binding of transcription factor and a reporter gene connected downstream from the promoter region; expression of the reporter gene is detected; and test DNA is isolated from a eukaryotic host in which expression has been detected;

(7) The method described in (6) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA binding domain, and a transcription activation domain;

(8) The method described in (6) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and a GAL4 transcription activation domain, and wherein the promoter region that is activated by binding of the transcription factor is the promoter region of a GAL1 gene in which the operator sequence has been replaced with the LexA operator sequence;

(9) The method described in (8) wherein the nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5;

(10) The method described in any of (6)-(9) wherein the

reporter gene is the LEU2 and/or β-galactosidase gene;

(11) A vector having an incorporation site of test DNA adjacent to DNA coding for transcription factor not having nuclear transportability;

(12) The vector described in (11) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA binding domain, and a transcription activation domain;

(13) The vector described in (11) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and the GAL4 transcription activation domain;

(14) The vector described in (13) wherein the nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5;

/6

(15) A kit comprising: (1) a vector having an incorporation site for test DNA adjacent to DNA coding for transcription factor not having nuclear transportability; and (2) a eukaryotic host having in its nucleus an expression unit comprising a promoter region activated by binding of the transcription factor and a reporter gene connected downstream from the promoter region;

(16) The kit described in (15) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA binding domain, and a transcription activation domain;

(17) The kit described in (15) wherein the transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and a GAL4 transcription activation domain; wherein the promoter region that is activated by binding of the transcription factor is the promoter region of a GAL1 gene in which the operator sequence has been replaced with the LexA operator sequence; and wherein the eukaryotic host is yeast;

(18) The kit described in (17) wherein the nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5; and

(19) The kit described in any of (15)-(18) wherein the reporter gene is the LEU2 and/or β-galactosidase gene.

In the present invention, the term "transcription factor" means a protein having a DNA binding domain and a transcription activation domain that activates the transcription of a specific gene, regardless of whether or not it occurs naturally. Further, in the present invention, the term "peptide" includes, in addition to proteins, partial peptides of proteins, synthetic

peptides, and the like.

The present invention relates first to a method of detecting the nuclear transportability of a peptide coded for by test DNA characterized in that fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is introduced into a eukaryotic host having in its nucleus a promoter region that is activated by binding of the transcription factor and a reporter gene the expression of which is induced by activation of the promoter region, and detecting expression of the reporter gene.

/7

In the present invention, the transcription factor employed in the preparation of "transcription factor not having nuclear transportability" is not specifically limited other than that it be capable of specifically controlling the expression of a gene in a eukaryotic organism; examples of transcription factor suitable for use are GAL4 (Giniger, E., Cell, 1985, 40: 767-774), p53 (Chumakov, P. M., Genetika, 1988, 24: 602-612), GCN4 (Hinnenbusch, A. G., Proc. Natl. Acad. Sci., 1984, 81: 6,442-6,446), VP16 (Triezeneberg, S. J., Genes. Dev., 1988, 2: 718-729), RelA (Nolan, G. P., Cell, 1991, 64: 961-969), Oct-1 (Strum, R. A., Genes. Dev., 1988, 2: 1,582-1,599), c-Myc (Watt, R., Nature, 1983, 303: 725-728), c-Jun (Angel, P., Cell, 1988, 55: 875-885), MyoD (Write, W. E., Cell, 1989, 56: 607-617), and the like.

The "transcription factor not having nuclear transportability" of the present invention is not particularly limited other than that it be transcription factor not having nuclear transportability (or having extremely low nuclear transportability) and having transcription activation ability and DNA binding ability. Examples are transcription factors in which the nuclear transport signal has been eliminated or replaced with other amino acids, transcription genes that are fused proteins comprising DNA binding domains and transcription activation regions, and the like.

Substances of low molecular weight (molecular weights not greater than 40,000 daltons) are generally thought to move by diffusion into nuclear holes other than by specific active transport systems. Even when the active nuclear transportability of transcription factor is eliminated due to loss or substitution of the nuclear transport signal, the movement of the substance into the nucleus by diffusion still occurs. In such cases, a signal can be added by localization of a protein in the cell outside the nucleus, thereby permitting complete or minimal control of substance movement into the nucleus by diffusion. The

"transcription factor not having nuclear transportability" of the present invention includes transcription factor to which is added in this manner a localized signal inside the cell but outside the nucleus. Examples of localized signals inside the cell but outside the nucleus are nuclear export signals (Gorlich, D., Science, 1996, 271: 1,513-1,518), secretion signals, peroxysome transport signals, rough-surfaced endoplasmic reticuli transport signals, mitochondria movement signals (Nakai, K., Genomics, 1992, 14: 897-911; Nakai, K., PSORT WWW server, http://psort.nibb.ac.jp/), and the like; the present invention is not limited thereto.

/8

Further, there are transcription factors having multiple nuclear transport signals and transcription factors in which it is entirely impossible to specify the position of a nuclear transport signal within the molecule despite the observation of nuclear transportability (GAL4, p. 53, and the like (TANAKA, Mahito, Cell Science (Japanese), 1991, 7: 265-272)). Further, there are transcription factors in which nuclear transport signals overlap with DNA binding domains or transcription activation domains, so that the loss or replacement of the nuclear transport signal may result in the loss of even DNA binding ability or transcription activation ability. When employing such transcription factors, even when it is impossible to completely specify the nuclear transport signal sequence, it suffices to specify the region required for eliminating nuclear transportability and remove or replace this region to prepare transcription regulating factor not having nuclear transportability. Further, an artificial hybrid transcription factor in which the DNA binding domain of a protein derived from a eukaryotic or prokaryotic organism that is known not to contain a nuclear transport signal and a transcription activation domain that is known not to contain a nuclear transport signal can be created to prepare transcription factor. The "transcription factor not having nuclear transportability" in the present invention includes transcription factor thus prepared.

The transcription activation domain employed in the preparation of the transcription factor not having nuclear transportability of the present invention includes, but is not limited to, the GAL4 transcription activation domain (Brent, R., Cell, 1985, 43: 729-736), Bicoid, c-Fos, c-Myc, v-Myc, B6, B7, B42 (Golemis, A. E., Mol. Cell. Biol., 1992, 12: 3,006-3,014), GCN4 (Hope, I. A., Cell, 1986, 46: 885-894), and VP16 (Clontech Co., Mammalian MATCHMAKER Two-Hybrid Assay Kit). The DNA binding domain includes, but is not limited to, GAL4 (Giniger, E., Cell,

1985, 40: 767-774), p53 (Chumakov, P. M., Genetika, 1988, 24: 602-612), GCN4 (Hinnenbusch, A. G., Proc. Natl. Acad. Sci., 1984, 81: 6,442-6,446), VP16 (Triezeneberg, S. J., Genes Dev., 1988, 2: 718-729), RelA (Nolan, G. P., Cell, 1991, 64: 961-969), Oct-1 (Strum, R. A., Genes. Dev., 1988, 2: 1,582-1,599), c-Myc (Watt, R., Nature, 1983, 303: 725-728), c-Jun (Angel, P., Cell, 1988, 55: 875-885), MyoD (Write, W. E., Cell, 1989, 56: 607-617), and other DNA binding domains that have been identified in transcription factor.

/9

"DNA coding for transcription factor not having nuclear transportability" can be prepared, for example, by the method of partially or completely removing the DNA sequence coding for nuclear transport signals from the DNA coding for transcription factor, the method of replacing the sequence within the nuclear transport signal by the incorporation of a site-specific variation, the method of adding a localized signal outside the nucleus but inside the cell, the method of fusing the transcription activation domain with the DNA binding domain, and suitable combinations of these methods. The general genetic operations in these methods are described in the literature (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

The "test DNA" employed in the method of the present invention includes cDNA, genome DNA, synthetic DNA, and the like that is not specifically limited beyond that it be DNA coding for a protein or a component peptide thereof. DNA coding for transcription factor not having nuclear transportability can be fused with test DNA by the usual methods (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

The fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is usually inserted into a suitable expression vector and introduced into an eukaryotic host. The expression vector is not particularly limited other than that it be capable of stably expressing a protein coded for by the fused DNA of DNA coding for specific transcription factor from which the nuclear transportability has been eliminated and test DNA; however, an expression vector functioning as a shuttle vector stably maintained by both the host and *E. coli* is preferred. For example, when baker's yeast is employed as

/10

the host, a unit for the expression of the protein (where the

expression unit comprises a promoter region functioning within yeast (the promoter region of ADH1, GAL1, or the like), an expressed protein code region, a multicloning site, and a terminator region (the terminator region ADH1 or the like)) can be incorporated for use into an embedded vector that is embedded in a yeast chromosome not having a replication starting point within the vector, a plasmid vector (centromea vector (low copy), 2μ vector (high copy), and the like are commercially available) present as a plasmid and having a replication starting point within the vector. Specifically, embedded vectors and centromea vectors are commercially available from Stratagene Co. as "pRS vectors" having various nutritional requirement marker genes (LEU2, HIS3, URA3, TRP1, and the like) for complementing the nutritional requirements of the host. Variant host strains corresponding to the respective marker genes are included as kits. Various commercially available vectors (Stratagene Co.'s HybriZapII, GAL4 Two-Hybrid Phagemid vector, Clontech Co.'s Matchmaker vector, and the like) employed in "Two-Hybrid systems" having nutritional requirement marker genes (LEU2, HIS3, URA3, TRP1, and the like) for complementing the nutritional requirements of the host, and the variant host strains corresponding to the respective vectors, can be employed as 2μ vectors. When employing animal cells as host, commercially available common mammal expression vectors in the form of vectors embedded in chromosomes (such as pMAM, pMAM-neo, and the like from Clontech Co.) or vectors maintained as episome (λDR2, pDR2 vector systems and the like from Clontech Co.) may be combined with suitable host animal cells (CHO, Mouse Fibroblast, Hela, U937, BHK, and the like) for use. The vector pMT2 and the like for transient expression using COS cells or the like (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) may also be employed. The insertion of the above-described fused DNA into the expression vector may be conducted by the usual methods (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Further, the "eukaryotic host" into which the above-described fused DNA is incorporated in the present invention is not specifically limited other than that it be a eukaryotic host capable of stably expressing proteins coded

/11

for by the above-described fused DNA. However, from the perspectives of convenience of handling, ease of incorporation and recovery of genes, safety, and the like, yeast and animal

cultured cells are particularly desirable. The eukaryotic host employed in the present invention has within its nucleus a promoter region that is activated by the binding of a specific transcription factor and a reporter gene connected downstream from this promoter region.

The "promoter region that is activated by the binding of a specific transcription factor" includes an upstream activating sequence (UAS) for binding of transcription factor or a cys control region called an operator sequence and a TATA sequence and is not particularly limited other than that it be a promoter region that is activated to specific transcription when transcription factor binds to the UAS. For example, in the case of baker's yeast, an example of a cys control region is natural GAL1 UAS (comprising four GAL4 binding sequences), artificial GAL1 UAS (comprising three GAL4 binding sequences), LexA UAS (comprising 1-8 LexA binding sequences)(Estojak, J., Mole. Cell. Biol., 1995, 15: 5,820-5,829). Further, examples of TATA sequences are GAL1 TATA, CYC1 (cytochrome C1) TATA, LEU2 TATA, and HIS3 TATA. These cys control regions and TATA sequences can be combined to construct various promoter regions of differing expression levels and inducement conditions (Clontech Co., Yeast Protocols Handbook, PT3024-1: 5-8). That is, a promoter region in which a transcription factor binding sequence is present in the cys control region and the activity of the promoter is controlled by the transcription factor suffices.

Further, in baker's yeast, the genetic analysis of which is quite advanced, the use as reporter gene of a gene relating to the nutritional requirements of the host (LEU2, HIS3, TRP1, URA3, or the like), a gene (such as GAL1) relating to the exploitation of required nutritional sources, or a gene compensating for the loss or damage of some other gene required for survival makes it possible to readily detect the expression of the gene through the survival or death of the host. It is also possible to employ a generally known reporter gene that can be detected by the activity of an enzyme such as β-galactosidase, chloramphenicol acetyltransferase, or luciferase, or green fluorescent protein (from Clontech Co.) permitting the

/12

direct detection of fluorescent light while [the transformants are] alive. Further, the above-described general-use reporter genes as well as drug-resistance genes may be employed in animal cells.

The above-described promoter region and reporter gene may be spliced by the usual methods (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory

Press, Cold Spring Harbor, NY).

For example, when employing baker's yeast as host, the usual methods, such as the lithium acetate method (Clontech Co., Yeast Protocols Handbook, PT3024-1: 17-20), can be used for the genetic introduction of the promoter region activated by the binding of transcription factor and the reporter gene that is spliced downstream from this promoter region. Based on differences in the vector employed (either the above-described embedded vector or a plasmid vector), the target gene may be selectively incorporated onto a chromosome or placed within the nucleus as a plasmid. Gene introduction is also possible by the usual methods, such as the ribosome method (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), in animal cells. Based on differences in the vector employed (either the above-described embedded vector or an episome vector), the target gene may be selectively incorporated onto a chromosome or placed within the nucleus as an episome.

Further, a commercially available eukaryotic host organism may be employed as the eukaryotic host organism having the above-described promoter region and reporter gene within the cell. For example, when employing LexA as the transcription factor DNA binding domain, the yeast EGY48[p8OP-lacZ] (available from Clontech Co.) characterized by comprising a promoter region having the LexA operator sequence, LEU2, which is the downstream reporter gene thereof, and β-galactosidase on the chromosome and on plasmids, respectively, can be employed.

For example, when employing baker's yeast as host, the vector containing the fused DNA of DNA coding for specific transcription factor the nuclear transportability of which has been eliminated and test DNA can be incorporated into the eukaryotic host by the usual methods, such as the lithium acetate method (Clontech Co., Yeast Protocols

/13

Handbook, PT3024-1: 17-20). Based on differences in the vector employed (either the above-described embedded vector or a plasmid vector), the target gene may be selectively incorporated onto a chromosome or placed within the nucleus as a plasmid. Gene introduction is also possible by the usual methods, such as the ribosome method (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), in animal cells. Based on differences in the vector employed (either the above-described embedded vector or an episome vector), the target gene may be selectively incorporated onto a chromosome or placed within the nucleus as an episome.

For example, in baker's yeast, the genetic analysis of which is quite advanced, the use as reporter gene of a gene relating to the nutritional requirements of the host (LEU2, HIS3, TRP1, URA3, or the like), a gene (such as GAL1) relating to the exploitation of required nutritional sources, or a gene compensating for the loss or damage of some other gene required for survival makes it possible to readily detect the expression of the reporter gene in the transformant thus obtained through the survival or death of the host. It is also possible to employ a generally known reporter gene that can be detected by the activity of an enzyme such as β-galactosidase, chloramphenicol acetyltransferase, or luciferase, or green fluorescent protein (from Clontech Co.) permitting the direct detection of fluorescent light emitted by living cells. Further, the above-described general-use reporter genes as well as drug-resistance genes may be employed in animal cells to detect expression. As a result, if expression of the reporter gene is detected, it may be concluded that the test DNA codes for a peptide having nuclear transportability, and if expression of the reporter gene is not detected, it may be concluded that the test DNA does not code for a peptide having nuclear transportability.

/14

Second, the present invention relates to a method of isolating test DNA coding for a peptide having nuclear transportability, characterized in that fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is introduced into a eukaryotic host having in its nucleus a promoter region that is activated by the binding of transcription factor and a reporter gene connected downstream from the promoter region; expression of the reporter gene is detected; and test DNA is isolated from a eukaryotic host in which expression has been detected. The test DNA can be isolated from a eukaryotic host in which expression of the reporter gene has been detected by, for example, in the case of baker's yeast when the test DNA is present on a plasmid (yeast-*E. coli* shuttle vector), refining plasmid from a single colony, using the plasmid obtained to transform *E. coli*, and further refining plasmid from the transformant. Alternatively, complete DNA from a single colony can be refined, and the refined DNA used as a template to amplify and refine the test DNA by PCR (Clontech Co., Yeast Protocols Handbook, PT3024-1: 29-37). As regards animal cells, as well, complete DNA is basically refined from a single colony and employed as a template to amplify and isolate the test DNA by PCR.

The present invention further relates to a kit comprising: a

vector having an incorporation site for test DNA adjacent to DNA coding for transcription factor not having nuclear transportability; and a eukaryotic host having in its nucleus an expression unit comprising the vector, a promoter region binding transcription factor, and a reporter gene spliced downstream from the promoter region. Test DNA is introduced at a test DNA introduction site in the vector of the present invention, and the vector is introduced into a eukaryotic host having in its nucleus an expression unit comprising a promoter region binding the transcription factor and a reporter gene spliced downstream from the promoter region. The test DNA introduction site is usually the only site on the vector that can be cleaved by a specific control enzyme. When expression of the reporter gene in the eukaryotic host is detected as a result of the introduction of the vector into the eukaryotic host, it is concluded that the test DNA that has been introduced into the vector codes for a peptide having nuclear transportability, and when expression of the reporter gene is not detected, it is concluded that the test DNA that has been introduced into the vector does not code for a peptide having nuclear transportability. Thus, it is readily possible to determine whether or not the test DNA codes for a peptide having nuclear transportability, and the isolation of DNA coding for peptides having nuclear transportability

/15

is readily accomplished. Specifically, a DNA library can be built with the above-described vectors, this library can be introduced into the above-described eukaryotic hosts, and the expression of reporter genes can be detected to efficiently and comprehensively isolate DNA coding for peptides having nuclear transportability from within the library.

Brief Description of the Figures

Fig. 1 shows the plasmid "pLexAD".

Fig. 2 shows the plasmid "pLexADrev".

Fig. 3 shows the plasmid "pRS1F".

Fig. 4 shows the plasmid "pRS3F".

Fig. 5 shows an assay for nuclear transportability in transcription factor by fusion with a test peptide.

Fig. 6 shows an assay for nuclear transportability in transcription factor fused to a test peptide.

Fig. 7 shows the plasmid "pNS".

Fig. 8 shows an assay for nuclear transportability in various peptides using the plasmid "pNS".

Best Mode of Implementing the Invention

Embodiments of the present invention are specifically described below; however, the present invention is not limited to these embodiments. In the embodiments set forth below, except where specifically stated otherwise, the basic genetic engineering methods employed were those described in the literature (Sambrook, J., Molecular Cloning: A Laboratory Manual, 1989, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Restriction enzymes, other modifying enzymes, and other genetic engineering products were purchased from Hōshuzō and the use conditions of the respective accompanying manuals were adhered to. Further, a "QIAprep Kit" (from Qiagen Co.) was employed to refine the plasmids from *E. coli*. An "ABI Prism 377" (from Perkin Elmer Co.) was employed to verify base sequences. Reagents

/16

from that company were employed to prepare samples for analysis and the methods employed conformed to the product manuals. Handling of the yeast (culture media, host, shuttle vectors, gene introduction methods, reporter gene assay method, gene isolation, and the like) was conducted with a "Matchmaker LexA Two-Hybrid System" (from Clontech Co.) according to the accompanying "Yeast Protocols Handbook". Synthesis of custom oligonucleotide was farmed out to Tōa Gōsei Co.

[Embodiment 1] Preparation of DNA sequences coding for GAL4 transcription activation domain by PCR

(1) Amplification by PCR of DNA sequences coding for the GAL4 transcription activation domain

A DNA fragment comprising the GAL4 transcription activation domain (the base sequence of which is given by sequence number 3) was amplified with the "GeneAmp PCR System 2400" (Perkin Elmer Co.) using a template in the form of "Plasmid pACT2" (Clontech Co.) and primers in the form of "Primer NU13" (sequence number 1) with an add-in EcoRI site designed into the 5' end and "Matchmaker 3' AD LD-Insert Screening Amplimer" (sequence number 2) (Clontech Co.). "TaKaRa Ex Taq" (TaKaRa Co.) was employed as the Taq polymerase and the product manual was adhered to for the reaction conditions and the like. The DNA fragment that had been amplified in this manner was refined by precipitation from ethanol and digested with the restriction enzymes EcoRI and NcoI. Six percent polyacrylamide gel electrophoresis was conducted and the targeted DNA fragment was cut out of the gel and recovered by electroelution.

(2) Preparation of the vector "pLexAD" expressing a fused protein of the LexA protein and GAL4 transcription activation

domain

The DNA fragment of (1) above coding for the GAL4 transcription activation domain was inserted between the EcoRI site and the NcoI site in the multicloning sites of the plasmid "pLexA" (Clontech Co.) to build "pLexAD" (Fig. 1). The base sequence was determined to verify that the targeted segment had indeed been inserted. The base sequence of the LexA gene is given by sequence number 4.

(3) Preparation of the vector "pLexADrev" in which the nuclear export signal (NES) is inserted on the N end of LexA

/17

A nuclear export signal (sequence number 5) having the Rev protein of HIV was synthesized in the following manner and inserted at the HpaI site near the N end of the LexA protein coded for by "pLexAD". "NU9" (sequence number 6) was synthesized as the sense chain and "NU10" (sequence number 7) as the antisense chain, the two were phosphorylated on the 5′ end with T4 polynucleotide kinase, and the two were annealed. The DNA fragment was inserted into "pLexAD" that had been predigested with HPaI and dephosphorylated with alkali phosphatase to construct "pLexADrev" (Fig. 2). The base sequence was determined to verify that the targeted segment had indeed been inserted.

(4) Construction of the plasmid "pRS1F" having a CEN/ARS region at a replication starting point for the expression of a fused protein of LexA protein and a GAL4 transcription activation domain and construction of the plasmid "PRS3F" having a CEN/ARS region at a replication starting point for the expression of a fused protein of LexA protein with an inserted NES and a GAL4 transcription activation domain

The minimum unit required for the expression in yeast of a fused protein of common LexA protein not having an inserted nuclear export signal (NES) and a GAL4 transcription activation domain, and a fused protein of LexA protein with an NES inserted at the N end and a GAL4 transcription activation domain (the base sequence in which the amino acid sequence of this fused protein is recorded in combined form is given in sequence number 8), is a DNA fragment of about 1.7 kb obtained by digesting "pLexAD" with SphI for the former, and digesting "pLexADrev" with SphI for the latter. This expression unit comprises an ADH1 promoter region, an expression protein coding region, a multicloning site, and an ADH1 terminator region. After refining the DNA fragments of these respective expression units, the portion of a PvuII digested fragment comprising in advance the multicloning site of the plasmid "pRS413" (Stratagene Co.) (yeast shuttle vector, CEN/ARS origin) was inserted at the SphI site of the vector

"pRSF" that had been substituted with PvuII digested fragment comprising the multicloning site of the widely used plasmid pUC19 to construct "pRS1F" and "pRS3F" (Figs. 3 and 4, respectively). The base sequence was determined to verify that the targeted segment had indeed been inserted. Since

/18

a pLexA-derived multicloning site was present immediately following the fused protein functioning as a transcription factor in the "pRS1F" (positive control) and "pRS3F" constructed in this manner, a targeted DNA fragment such as cDNA could be readily fused by the usual methods and expressed.

[Embodiment 2] Validation of the effectiveness of the nuclear transport protein trap vector "pRS3F" by fusion of artificial nuclear transport protein cDNA

(1) Fusion of a known cDNA fragment

A cDNA fragment coding for the branch strand amino acid binding protein ('BraC) of *Pseudomonas aeruginosa* from which the secretion signal observed to be locally present in the cytoplasm had been removed (a base sequence in which the amino acid sequence of this protein is recorded in combination is shown in sequence number 9) (TANAKA, Mahito, New Biochemistry Experiment Lecture 6 (Ed. by the Japan Biochemistry Society), Biomembranes and Membrane Transport (2/2), 1992, Tokyo Chemistry Club, 9 15) and an artificial nuclear transport protein with SV40 large T antigen-derived nuclear transport signal fused onto its N end was fused in-frame onto the C end of the GAL4 transcription activation domain of "pRS3F" as a known cDNA fragment. More precisely, "pRS3F'BraC" was constructed by inserting the DNA fragment (NcoI-DraI) coding for "'BraC" into "pRS3F" that had been refined by digestion with XhoI, Klenow treatment to smooth off the ends, and digestion with NcoI. "pRS3FN'BraC" was then constructed by inserting a synthetic DNA fragment coding for a nuclear transport signal (sequence number 10) derived from SV40 large T antigen, that is, synthesizing "NU17" as sense strand (sequence number 11) and "NU18" as antisense strand (sequence number 12), phosphorylating the 5' ends thereof with T4 polynucleotide kinase, and annealing the two, into a vector obtained by refining this "pRS3F'BraC" by digestion with NheI and NcoI. Further, as a control test, "pRS3FN" having only a nuclear transport signal and no "'BraC" fragment was constructed in the same manner. Correct insertion of the targeted fragment was confirmed by determining the base sequence.

(2) Nuclear transport capability assay based on reporter gene expression]

/19

The three plasmids "pRS3F'BraC", "pRS3FN'BraC", and pRS3FN" described in (1) above and the "pRS1F" and "pRS3F" constructed in Embodiment 1 were used to transform host yeast EGY48[p80P-lacZ] (obtained from Clontech Co.) having a promoter region (sequence number 13) having the LexA operator sequence (Estojak, J., Mole. Cell. Biol., 1995, 15: 5,820-5,829), the LEU2 reporter gene downstream therefrom, and β-galactosidase on chromosomes and on plasmid, respectively. Introduction into the host of the targeted plasmids was confirmed by complementation of HIS, a nutritional requirement marker. Next, the respective transformants were replicated in culture media (SD/-LEU, -HIS, -URA, X-gal) to assay expression of the reporter gene and cultured for 2-3 days at 30°C. As a result, both the reporter gene β-galactosidase and LEU2 were expressed, and blue coloration and normal development were confirmed, in the transformants into which had been introduced "pRS3FN'BraC" fused with artificial nuclear transport protein, "pRS3FN" fused with only nuclear transport signal, and "pRS1F" as a positive control (Figs. 5 and 6). By contrast, almost no reporter gene was expressed and neither blue coloration nor growth was observed in the transformants into which had been introduced "pRS3F'BraC" fused with a protein having no nuclear transport signal and "pRS3F" that had not been fused with anything (Figs. 5 and 6).

From these results, the in-frame fusion of a DNA fragment coding for a certain peptide onto the C end of transcription factor coding for "pRS3F" and the expression thereof in yeast permitted the detection of the presence or absence of nuclear transportability by using the expression of the reporter gene as indicator.

[Embodiment 3] Construction of the vector pNS for creating a cDNA library

"pRS3F" was improved. The improvement consisted of the following three points: (1) elimination of the EcoRI sites in the LexA and GAL4AD binding portions, (2) the introduction of an EcoRI site at the multicloning site, and (3) the elimination of unneeded regions derived from "pRS413" to achieve the smallest size possible.

First, a synthesis linker, "NU31" as sense strand (sequence number 14), and "NU30" as antisense strand (sequence number 15), were inserted at the EcoRI site of "pLexADrev" to obtain the plasmid "pLexADrev-dE".

/20

A DNA fragment comprising an about 1.7 kb ADH1 expression unit obtained by digestion of "pLexADrev-dE" with the restriction enzyme SphI was subcloned at the SphI site of the widely employed

plasmid pUC19 to obtain the plasmid "pULexADrev-dE". Next, a synthesis linker having an EcoRI site, a sense strand in the form of "NU28" (sequence number 16), and an antisense strand in the form of "NU29" (sequence number 17) were inserted between the NheI site and the NcoI site of the "pULexADrev-dE" to obtain the plasmid "pULexADrev-E". Further, "pRS413" was digested with DraIII and PvuII to remove a DNA fragment comprising a 757 bp multicloning site, and a synthesis linker having an SpHI site, a sense strand in the form of "NU25" (sequence number 18), and an antisense strand in the form of "NU26" (sequence number 19) were inserted at the removal site to obtain the plasmid "pRS-S". A DNA fragment comprising an about 1.7 kb ADH1 expression unit obtained by digesting the above-described "pULexADrev-E" with SphI was inserted at the SphI site of the "pRS-S" to construct the vector pNS (fig. 7) for use in creating a cDNA library (the transcription direction of ADH1 is identical to that of HIS3).

[Embodiment 4] Creation of a fused protein expression library (derived from precursor cells of the cultured human cell NT2) and a nuclear transport assay

(1) Creation of a fused protein expression library

mRNA was prepared by culturing precursor cells (Stratagene Co.) of the cultured human cell NT2 according to the supplemental protocol (Catalog #204101, Revision #036002a) and using the a commercial total RNA extraction kit and an mRNA extraction kit (Pharmacia Co.). Using a portion thereof (3 µg), a cDNA library was created using a commercial cDNA synthesis kit (Pharmacia Co.). Specifically, cDNA synthesis was conducted using oligo(dT)12-18 primer and inserted at the EcoRI/NotI site of the pNS vector. The cDNA was unidirectionally introduced using a Directional Cloning Toolbox (Pharmacia Co.). Subsequently, a portion of the cDNA library was employed to transform commercial *E. coli* (ElectroMAX DH10B Cells from GibcoBRL Co.) by electroporation (Gene Pulser from BIO RAD Co.) conducted in the usual manner (New Cytoengineering Test Protocol, Hidemasu Co., 114-115). The transformants

/21

obtained were cultured for 16 hr at 30°C in an LB agar medium comprising ampicillin (100 µm/mL), and after collecting the bacteria, plasmid was prepared (Qiagen Maxi kit from Qiagen Co.).

(2) Nuclear transport assay employing yeast

Using 60 µg of the plasmid of the fused protein expression library that had been prepared, EGY48 strain was transformed by the usual methods (Clontech Co., Yeast Protocols Handbook, PT3024-1: 17-20). When the transformants were cultured for 3-7 days at 30°C in SD agar medium (-His/-Leu) to select the clones

based on expression of the reporter gene LEU2, about 1,000 positive clones were obtained.

(3) Determination of base sequences

The base sequences of the cDNA fragments inserted into the vector were determined for some (12) of the positive clones thus obtained. To determine the base sequence, colony PCR was first employed to prepare template DNA from each of the clones. A small quantity of bacteria scraped from each clone was added to 20 µL of PCR reaction solution (0.5 unit of heat-resistant DNA polymerase (Ex Taq from TaKaRa Co.), 4 nmol of dNTP mixture, 0.4 pmol each of "primer NU15" (sequence number 20) and "primer NU36" (sequence number 21), 2 µL of supplemental buffer, and sterilized water) and the inserted cDNA fragments were subjected to 40 cycles of amplification using a "GeneAmp PCR System 2400" (Perkin Elmer Co.) at a denaturation [temperature] of 94°C, an annealing [temperature] of 60°C, and an expansion [temperature] of 72°C. Each PCR product was subjected to desalting with a Microcon-100 (Millipore Co.) and unreacted primer was removed to obtain template DNA. A portion (100-200 ng) of the template thus obtained was used to determine the base sequence by the method described in a product manual from ABI Co.

(4) Database analysis of the clones obtained

The base sequence of each clone was searched for in the Basic BLAST (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform=0) of the National Center for Biotechnology Information (NCBI), a public database. As a result, all 12 clones matched previously known genes. Of those, there were reports or suggestions that ten might function within the

/22

nucleus. Of the ten, five of the clones were: NP220 having nuclear transport signal-like sequences of the SV40 large T antigen type, rich in basic amino acids (Inagaki, H., J. Biol. Chem., 1996, 271: 12,525-12,531), PC4 (Ge, H., Cell, 1994, 78: 513-523), ERC-55 (Imai, T., Biochem. Biophys. Res. Commun., 1997, 233: 765-769), histone-binding protein (O'Rand, M. G., Dev. Biol., 1992, 154: 37-44), and prothymocin α1 (Manrow, R. E., J. Biol. Chem., 1991, 266: 3,916-3,924). One clone was hnRNPA1, which has an M9 sequence performing round trip movement into and out of the nucleus (Michael, W. M., Cell, 1995, 83: 415-422). Four more of the clones were ferritin H chain not having known nuclear transport signals (Cai, C. X., J. Biol. Chem. 1997, 272: 12,831-12,839), Shaperonin 10 (Bonardi, M. A., Biochem. Biophys. Res. Commun., 1995, 206: 260-265), protein kinase C inhibitor-I (Brzoska, P. M., Proc. Natl. Acad. Sci., 1995, 92: 7,824-7,828), and steroid receptor coactivater-1 (Onate, S. A., Science, 1995,

270: 1,354-1,357). No known nuclear transport signal was found in the two remaining clones, for which no function within the nucleus has yet been reported: tropomyocin (Lin, C. S., Mol. Cell. Biol., 1988, 8: 160-168) and G-rich sequence factor-1 (Qian, Z., Nucleic Acids Res., 1994, 22: 2,334-2,343);

[Embodiment 5] Creation of a fused protein expression library (derived from human fetal brain [cells]) and a nuclear transport assay

(1) First, a commercial human fetal brain cDNA library (Superscript library from GibcoBRL Co.) was amplified according to the protocol provided by the manufacturer. Plasmids comprising the cDNA fragments as inserts were then prepared using a plasmid manufacturing kit from Qiagen Co. Next, cDNA fragments cut out from a portion (30 µg) thereof using the two restriction enzymes EcoRI and NotI were sorted to obtain cDNA 0.7-4 kb in length by 0.8 percent agarose electrophoresis. The cDNA fragments thus obtained were inserted at the EcoRI/NotI site of the above-described pNS

/23

vector to prepare a fused protein expression library. A portion thereof was employed to transform commercial *E. coli* (ElectroMAX DH10B Cells from GibcoBRL Co.) by electroporation (Gene Pulser from Biorad) using the usual method (New Cytoengineering Test Protocol, Hidemasu Co., 114-115). The transformants obtained were cultured for 16 hr at 30°C in an LB agar medium comprising ampicillin (100 µm/mL), the bacteria were collected, and plasmid was prepared (Qiagen Maxi Kit from Qiagen).

(2) Nuclear Transport Assay Employing Yeast

EGY48 strain was transformed by the usual method (Clontech Co., Yeast Protocols Handbook, PT3024-1: 17-20) using 60 µg of plasmid from the fused protein expression library that had been prepared. When cultivated for 3-7 days at 30°C in SD agar medium (-His/-Leu) to select clones based on expression of the reporter gene LEU2, about 1,000 positive clones were obtained.

(3) Base sequencing$

The base sequences of the cDNA fragments inserted into the vector were determined for some (489 clones) of the positive clones thus obtained. To conduct sequencing, template DNA was prepared from each clone by colony PCR. A small quantity of bacterial matter scraped from each clone was added to 20 mL of PCR reaction solution (0.5 unit of heat-resistant DNA polymerase (Ex Taq from TaKaRa Co.), 4 nmol of dNTP mixture each, 0.4 pmol each of "primer NU15" (sequence number 22) and "primer NU36" (sequence number 23), 2 µL of supplemental buffer, and sterilized water) and the inserted cDNA fragments were subjected

to 40 cycles of amplification using a "GeneAmp PCR System 9600" (Perkin Elmer Co.) at a denaturation [temperature] of 94°C, an annealing [temperature] of 60°C, and an expansion [temperature] of 72°C. Each PCR product was subjected to desalting with a Microcon-100 (Millipore Co.) and unreacted primer was removed to obtain template DNA. A portion (100-200 ng) of the template thus obtained was used to determine the base sequence by the method described in a product manual from ABI Co.

(4) Database analysis of the clones obtained

/24

The base sequence of each of the 489 clones was searched for in the Basic BLAST (http://www.ncbi.nlm.nih. gov/cgi-bin/BLAST/nph-blast?Jform=0) of the National Center for Biotechnology Information (NCBI), a public database. As a result, 250 of the clones matched genes coding for 97 known proteins (Tables 1 and 2), 220 of the clones were either new sequences that were candidates for genes coding for new nuclear transport proteins or matched 172 genes coding for known expressed sequence tags (EST). Another 19 of the clones were either derived from nontranslation regions of known genes or had shifted codon read frames.

Table 1 shows those of the genes isolated by the method of the present invention that code for proteins that have been reported to have functions within the nucleus, and Table 2 shows those for which no function within the nucleus has been reported.

Table 1

Gene *	GenBank Accession	Function	Starting position of region where obtained*	Structural character-istics of region where obtained	Length (kb) of region where obtained	Medline Ul[4]
1		RNA binding protein				
2		Synapse/nuclear protein				
3		Saccharolytic enzyme				
4		Bacteria [illeg.]/Signal transmitting gene				
5		Transcription factor				
6		Calcium binding protein				
7		Transcription factor				
8		Cyclosporin binding protein				
9		Steroid receptor conjugate factor				
10		Transcription factor				
11		Gu and p53 interaction nuclear protein				
12		Centromere region interaction protein				
13		Transcription regulating factor				
14		Transcription factor				
15		DNA cleaving/modifying complex				

16		Transcription regulating factor				
17		Ribonucleic protein				
18		Ribonucleic protein				
19		Heterochromatin protein				
20		Transcription regulating factor				
21		NLS dependent nuclear transport receptor				
22		NLS dependent nuclear transport receptor				
23		Nuclear autoantigen				
24		DNA binding protein				
25		Metabolic enzyme				
26		Assumed transcription regulating factor				
27		M-phase phosphoprotein				
28		Major nuclear matrix protein				
29		Transcription factor				
30		DNA binding protein				
31		DNA binding protein				
32		Contributes to inducing cell proliferation				
33		Transcription regulating protein				
34		Protein interacting with homeotic protein BM11				
35		Assumed transcription regulating factor				
36		Protein interacting with kinesin-related proteins				
37		US snRNP subunit protein complex				
38		Transcription regulating factor				
39		Assumed transcription regulating factor				
40		Transcription factor				
41		Transcription factor				
42		Transcription factor				
43		DNA cleaving/modifying enzyme				
44		Nuclear membrane complex interacting protein				
45		Assumed transcription regulating factor				

/25

Table 2

Gene *	GenBank Accession	Function	Starting position of region where obtained*	Structural characteristics of region where obtained	Length (kb) of region where obtained	Medline Ul[4]
1		Contributes to purine synthesis path				
2		Saccharolytic enzyme				
3		Actin binding protein				
4		Pituitary protein				
5		MAPKKK Mammal homolog				

6		Similar to ME491/CD63 superfamily				
7		Contributes to intracellular protein transport				
8		Assumed colorectal cancer suppressing gene product				
9		Actin binding protein				
10		Metabolic enzyme				
11		Metabolic enzyme				
12		Contributes to signal transmission				
13		Endoplasmic reticulum calcium binding protein				
14		Actin binding protein				
15		Intermediate filament				
16		Testes/brain specific GST				
17		Assumed Golgi complex protein				
18		Similar to yeast CDC10				
19		Cyclin G interacting kinase				
20		Homolog of drosophila sina				
21		Function unknown				
22		Function unknown				
23		Function unknown				
24		Function unknown				
25		Function unknown				
26		Function unknown				
27		Function unknown				
28		Function unknown				
29		Assumed kinesin receptor				
30		Kinesin motor protein super family				
31		G alpha 2 interacting protein				
32		Metabolic enzyme				
33		Kinesin motor protein super family				
34		Contributes to [illeg.] transport				
35		Similar to nel protein				
36		aglycon sugar protein family				
37		Intermediate filament				
38		Actin binding protein				
39		Phosphoglycarate mutase family				
40		RacI interacting				

		protein				
41		Effector protein of small GTPase Rab5				
42		Small GTPase Rab5 interacting protein				
43		Intermediate filament interacting protein				
44		Ribosome protein				
45		Assumed transcription factor				
46		Contributes to [illeg.] transport				
47		Similar to Grb-2 having an SH3 domain				
48		Signal transmission adapter molecule				
49		Assumed transcription control factor				
50		Contributes to thyroid cancer				
51		Cell adhesion factor				
52		Intermediate filament				

	遺伝子*	GenBank Accession	機能	取得領域の開始部位*	取得領域の構造的特徴	[illegible]の長さ(kb)	Medline UI*
1	9G8 splicing factor	L22253	RNA 結合タンパク質	FOSPSRSRSR→	S/R rich	1.8	94283389
2	AD amyloid NACP (synuclein)	L08850	シナプス／核 タンパク質	ILEDMPVDPD→		0.7	88316381
3	aldolase A	X05236	解糖酵素	*		1.8	92182008
4	beta catenin	Z19054	細胞接着／シグナル伝達分子	VELTSSLFRT→		1.9	97047308
5	c-fos	V01512	転写因子	*	NLS, bZIP	2.0	83221560
6	calmodulin	D45887	カルシウム結合タンパク質	PTEAELQDMI→		1.4	96114780
7	CREB-2	M86842	転写因子	GLVSPSNNSK→	NLS, bZIP	1.8	92279218
8	cyclophilin A	Y00052	サイクロスポリン結合タンパク質	*		0.8	95394146
9	F-SRC-1	U59302	ステロイドレセプター共役因子	AINQSKSEDQ→		2.5	96291002
10	GADD153 (CHOP)	S40708	転写因子	*	NLS, ZIP	1.0	93015930
11	Gu binding protein	U78524	Gu および p53 相互作用核タンパク質	LKQMVMSLRV→		2.4	97320420
12	hCENP-B	X55039	動原体領域相互作用タンパク質	EDEDDDDEE→		1.7	91372020
13	hCREM-2	D14826	転写調節因子	*	NLS, bZIP	2.0	94266757
14	heat shock factor 1 (TCF5)	M64673	転写因子	LEHVHGSGPY→	ZIP	1.9	91334376
15	HHR23A protein	D21235	DNA切断／補修複合体	IPGSPEPEHG→		1.8	98292259
16	HIRA	X77633	転写調節因子	GDFSTAFFNS→	NLS	2.4	95359996
17	hnRNPC	M16342	リボ核タンパク質	*		1.9	87257872
18	hnRNPK	S74678	リボ核タンパク質	YDPNFYDETY→	KNS	1.5	97381839
19	HP1Hs-gamma	U26312	ヘテロクロマチンタンパク質	KKKRDAADKP→	NLS	1.8	96278941
20	hSNF2b	D26156	転写調節因子	VEEKILAAAK→	NLS	0.7	94768902
21	importin alpha 3	U93240	NLS依存的核移行レセプター	ICLSAVQAAR→	arm	2.4	96270582
22	karyopherin alpha 3	D89618	NLS依存的核移行レセプター	SAQTQAVVQS→	arm	1.9	96270582
23	Ki nuclear autoantigen	U11292	核自己抗原	*	NLS	2.4	86141726
24	Ku protein p70	M32865	DNA結合タンパク質	DSFENPVLQQ→	NLS	0.8	89174787
25	lactate dehydrogenase	Y00711	代謝酵素	NKITVVGVGQ→		2.0	87053963
26	leucine zipper protein (hDIP)	Z50781	推定の転写調節因子	*	ZIP	2.0	97136879
27	M-phase phosphoprotein (mpp6)	X98263	M期りん酸化タンパク質	KKIISEEHWY→	NLS	1.7	97039687
28	matrin 3	M63483	主要核マトリックスタンパク質	DGQSDENKDD→	NLS	1.9	91236771
29	NF-kappa-B p65 subunit	M62399	転写因子	DDRHRIEEKR→	NLS	2.0	91173312
30	NP220	D83032	DNA結合タンパク質	IPTGDEKTVD→	NLS	2.5	96218178
31	nucleobindin precursor	M96824	DNA結合タンパク質	QAQLKEVWEE→	ZIP, EF-hand	1.9	82392352
32	nucleosome assembly protein (NAP)	M86667	細胞増殖調節に関与	IPEFWLTVFK→	NLS	2.0	94128073
33	PC4	U12979	転写調節因子	*	NLS	2.0	94340740
34	polyhomeotic 1 homolog (HPH1)	U89277	ホメオティックタンパク質BMI1相互作用タンパク	HGERDLGNPN→		2.4	97220024
35	RBP2=retinoblastoma binding protein	S66431	推定の転写調節因子	LLEVSLDETQ→	NLS	1.9	94020841
36	SMAP	U59919	キネシン関連タンパク質相互作用タンパク質	GLKHLMKRAL→	arm	1.9	98175913
37	spliceome associated protein (SAP 145)	U41371	U2 snRNPサブユニットタンパク質複合体	ETRLKEKKPG→	NLS	1.7	96154048
38	SWI/SNF complex subunit (BAF170)	U66616	転写調節因子	RYDFQNPSRM→	NLS, ZIP	2.5	96397413
39	tat interactive protein (TIP60)	U74667	推定の転写調節因子	*	NLS	1.9	96182937
40	TEF=thyrotroph embrionic factor	U44059	転写因子	YMDLDEFLLE→	NLS, bZIP	2.5	96219638
41	TFE3	X51330	転写因子	IQELELQAQI→	NLS	2.0	90249724
42	TFEB	M33782	転写因子	ELTDAESRAL→	NLS, bHLHZIP	1.9	90318407
43	topoisomerase IIb	U54831	DNA 切断／修復酵素	DADDDDDDNN→	NLS	1.8	98122794
44	TPR	U69668	核膜複合体相互作用タンパク質	IQMTRRQSVG→		2.8	97177132
45	TSC-22	U35048	推定の転写調節因子	MYAVREEVEV→	ZIP	1.8	96244585

表2

	遺伝子	GenBank Accession	機能	取得領域の開始部位	取得領域の構造的特徴	取得領域の長さ(kb)
1	ADE2H1	X53793	プリン生合成経路に関与	*		2.0
2	aldolase C	X07292	解糖酵素	YPALSAEQKK→		1.8
3	alpha-actinin	X15804	アクチン結合タンパク質	EQVEKGYEEW→	coiled-coil, EF-hand	2.5
4	antisecretory factor-1	U24704	下垂体タンパク質	*	NLS	1.8
5	ASK1	D84476	MAPKKK 哺乳類ホモログ	IRTLFLGIPD→		2.0
6	cell surface glycoprotein	D10653	ME491/CD63 スーパーファミリーに類似	*		2.0
7	coatomer protein (COPA)	U24105	細胞内タンパク質輸送に関与	GHYQNALYLG→	WD-40	2.3
8	colorectal mutant cancer protein	M62397	推定の大腸癌抑制遺伝子産物	EISSIGVSSS→	NLS	4.2
9	cytoskeletal tropomyosin TM30	X04588	アクチン結合タンパク質	*	coiled-coil	2.0
10	cytosolic malate dehydrogenase	D55654	代謝酵素	*		1.8
11	dihydrolipoamide dehydrogenase	J03620	代謝酵素	IPVNTRFQTK→		1.9
12	epsilon 14-3-3 protein	U28936	シグナル伝達に関与	QHVETELKLI→		1.8
13	ERC-55	X78669	小胞体カルシウム結合タンパク質	LKDKKRFEKA→	NLS, EF-hand	1.9
14	fibroblast tropomyosin TM30	X05276	アクチン結合タンパク質	YEEEIKLLSD→	coiled-coil	1.8
15	glial fibrillary acidic protein (GFAP)	J04569	中間フィラメント	QYEAMASSNM→	coiled-coil	2.4
16	glutathione S-transferase M3 (GSTM3)	J05459	精巣／脳特異的GST	ESSMVLGYWD→		0.7
17	golgin-95	L06147	推定のゴルジ複合体タンパク質	QYVAAYQQLT→	coiled-coil	2.5
18	hCDC10=CDC10 homolog	S72008	酵母のCDC10に類似	REHVAKMKKM→	NLS, ZIP	1.8
19	HsGAK	D88435	サイクリンG相互作用キナーゼ	QGPPEDLLSE→		2.0
20	hSIAH2	U76248	ショウジョウバエ sina のホモログ	EHEDICEYRP→		1.9
21	KIAA0116	D29958	機能未知	VTLSEAEKVY→		0.9
22	KIAA0136	D50926	機能未知	QLLLVTEEKE→		2.5
23	KIAA0171	D79993	機能未知	QATMTSSQSM→		2.2
24	KIAA0181	D80003	機能未知	QNLVSKETST→	NLS	0.8
25	KIAA0332	AB002330	機能未知	IPIDATPIDD→	NLS	1.7
26	KIAA0365	AB002363	機能未知	SGCPLQVKKA→	NLS	2.0
27	KIAA0373	AB002371	機能未知	IISATSQKEA→	NLS	0.6
28	KIAA0432	AB007892	機能未知	SAPIINFSAQ→		2.3
29	kinectin	L25616	推定のカイネシンレセプター	DKLQALANEQ→	coiled-coil	2.8
30	kinesin-2 (HK2)	Y08319	カイネシンモータータンパク質スーパーファミリー	RVKELTVDPT→	coiled-coil	2.4
31	LGN protein	U54999	G alpha i2 相互作用タンパク質	IPNSQRKISA→		1.9
32	malate dehydrogenase	U20352	代謝酵素	*		1.8
33	mitotic kinesin-like protein 1	X67155	カイネシンモータータンパク質スーパーファミリー	EFDNAVLSKE→	NLS, coiled-coil	2.6
34	N-ethylmaleimide-sensitive factor	U03985	膜輸送に関与	LASLENDIKP→		2.5
35	nel-related protein (NRP1)	D83017	ネルタンパク質に類似	RNQKHGLFKG→	EGF-like	2.3
36	neurocan (CSPG3)	AF026547	アグリカン類タンパク質ファミリー	PAQVNKAEHS→	NLS	2.0
37	neurofilament-66	S78296	中間フィラメント	LAFVRQVHDE→	coiled-coil	2.5
38	non-muscle myosin heavy chain-B	M69181	アクチン結合タンパク質	KKLKSLEAEI→	coiled-coil	2.5
39	phosphoglycerate mutase (PGAM-B)	J04173	ホスホグリセリン酸ムターゼファミリー	*		2.4
40	por1	X97567	Rac1相互作用タンパク質	FGRGSRRTVD→	coiled-coil	1.8
41	Rabaptin-5	X91141	small GTPase Rab5のエフェクタータンパク質	IQIQEAETRD→	coiled-coil	2.0
42	Rap2 interacting protein 8 (RPIP8)	U93871	small GTPase Rap2相互作用タンパク質	KFRIVYAQKG→		1.8
43	restin	X64838	中間フィラメント相互作用タンパク質	KFIKQADEEK→	NLS, coiled-coil	2.3
44	RIG-like 7-1	AF034208	リボソームタンパク質	DNQRDCQPGL→		1.8
45	RING zinc finger protein (RZF)	AF037204	推定の転写因子	KTKKTCPVCK→		1.9
46	secretogranin I(chromogranin B)	Y00064	膜輸送に関与	PEYGEEIKGY→	NLS	1.7
47	SH3GL2	X99657	SH3ドメインを持つGrb-2に類似	LHDKQLREIQ→	NLS, SH3	2.1
48	STAM	U43899	シグナル伝達アダプター分子	QPNWKGETH→	ITAM	2.4
49	tax1-binding protein TXBP151	U33821	推定の転写制御因子	SKEDTCFLKE→		1.9
50	TFG protein	Y07968	甲状腺癌に関与	LRRELIELRN→	coiled-coil	1.9
51	trophinin	U04811	細胞接着因子	PSNSIGFGAA→		1.9
52	vimentin	Z19554	中間フィラメント	ELQAQIQEQH→	coiled-coil	1.7

The symbols in the tables have the following meanings:

[a]: Indicates that the protein obtained represents the shortest inserted fragment in the group to which assigned.

[b]: Indicates 10 amino acid residues from the amino terminus of the protein coded for by the inserted gene fragment.

[c]: Medline Unique Identifier of the document reporting a function in the nucleus.

*: A clone comprising the entire translation region.

S/R rich: A serin/alginin rich region.

NLS: Assumed nuclear transport signal rich in basic residues.

/26

ZIP: Leucin zipper

bZIP: Basic leucin zipper

KNS: hnRNP K nuclear transport signal

arm: Armaggio repeat

bHLHZIP: Basic helix loop helix leucin zipper

SH3: Src homology domain 3

ITAM: Immunoreceptor tyrosine-based activation motif

As shown in Tables 1 and 2, about half of the 97 known proteins were proteins reported to have functions within the nucleus. The ratio of transcription control factors and DNA/RNA splicing proteins was particularly high. Accordingly, even as regards the new genes, it may be readily anticipated that genes coding for unknown proteins functioning within the nucleus will be efficiently and specifically obtained. Further, with regard to the hnRNPK protein among the isolated clones, the KNS sequence (Matthew, W., EMBO J., 1997, 16: 3,587-3,598) responsible for back and forth movement into and out of the nucleus was found. The finding of the M9 sequence and the KNS sequence that are responsible for movement into and out of the nucleus among the clones isolated by the method of the present invention demonstrates that the method of the present invention is not only capable of specifically selecting with high efficiency just nuclear transport proteins, but can also be expanded to the general selection of proteins moving into and out of the nucleus (outside the nucleus -> inside the nucleus, inside the nucleus -> outside the nucleus).

[Embodiment 6] Demonstration of the efficacy of the nuclear transport protein trap vector "pNS" based on the fusion of cDNA coding for known nuclear transport proteins

(1) Construction of fused plasmids of known cDNA fragments

cDNA in the form of "'BraC" (TANAKA, Mahito, New Biochemistry Experiment Lecture 6 (Ed. by the Japan Biochemistry

Society), Biomembranes and Membrane Transport (2/2), 1992, Tokyo Chemistry Club, 9 15) and calcium/calmodulin dependent protein kinase kinase "CaMKK" (Tokumitsu, H., J. Biol. Chem., 1995, 270 (33): 19,320-19,324; Tokumitsu, Hiroshi, "Localization of CaMKK in Cells", unreleased data) were employed as representative proteins localized in cytoplasm. cDNA in the form of SV40 "NLS", "NLS-'BraC" obtained by artificially fusing SV40 "NLS" and "'BraC", the transcription factor NF-kappa-B

/27

p65 subunit "NFKBp65" (Ganchi, P. A., Mol. Biol. Cell, 1992, 3(12): 1,339-1,352), and the transcription factor "c-Fos" (Tratner, I., Oncogene, 1991, 6(11): 2,049-2,053) was employed as representative proteins localized in the nucleus and having conventional nuclear transport signals. The plasmid "pRS1F" was employed for "LexAD", "pNS" for "NES-LexAD", "pRS3FN" for "NES-LexAD-NLS", "pRS3F'BraC" for "NES-LexAD-'BraC", and "pRS3FN'BraC" for "NES-LexAD-NFKBp65", respectively. "NES-LexAD-NFKBp65" was prepared by amplifying "NFKBp65" by PCR employing the primers "NU32" (sequence number 24) and "NU24" (sequence number 25) and employing "PME18S(N)-p65" (Tsuboi, A., Biochem. Biophys. Res. Commun., 1994, 199(2): 1,064-1,072) as template, refining the fragments by digestion with the restriction enzymes MunI and NotI, and inserting the fragments into the EcoRi/NotI site of "pNS". Similarly, "NES-LexAD-cFOS" was prepared by amplifying "c-FOS" by PCR employing the primers "NU34" (sequence number 26) and "NU24" and employing "PME18S(N)-cFos" (Tsuboi, A., Biochem. Biophys. Res. Commun., 1994, 199(2): 1,064-1,072) as template followed by insertion into the EcoRI/NotI site of "pNS". "NES-LexAD-CaMKK" was prepared by digesting "pET-CaMKK" (provided by Mr. Hiroshi TOKUMITSU) with the restriction enzyme NcoI to obtain a "CaMKK" cDNA fragment, which was then inserted at the NcoI site of "pNS".

(2) These plasmids were each introduced into EGY48 strain and expression of the reporter gene LEU2 was observed. Following transformation with the various plasmids described above in (1), direct plating on SD culture (-HIS, -LEU) was conducted. "LexAD" not having NES was thought to form colonies because of passive diffusion into the nucleus. The formation of colonies in "NES-LexAD" into which NES had been introduced was completely inhibited. However, in "NES-LexAD-NLS" into which NLS had been additionally incorporated, colony formation was again observed. Similarly, in "NES-LexAD-NLS-'BraC", "NES-LexAD-NFKBp65", and "NES-LexAD-cFos", all of which comprised conventional NES, colony formation was observed. In "NES-LexAD-'BraC" and "NES-LexAD-CaMKK" which did not have nuclear transportability, colony

formation was completely inhibited. These results demonstrate that the specific detection of cDNA fragments having nuclear transportability is possible in systems employing "pNS" vector.

/28

Potential For Industrial Use

Based on the present invention, it is possible to conveniently detect whether or not a peptide coded for by test DNA has nuclear transportability by employing as indicator the expression of a reporter gene. Further, it is possible to rapidly, efficiently, and comprehensively clone DNA coding for a protein having nuclear transportability by employing as indicator the expression of a reporter gene. Based on the present invention, not only is the obtaining of DNA coding for new intranuclear proteins of biological importance advanced, but extremely useful gene expression information (time, place, expression frequency, and the like) with regard to research into the functioning of proteins in the nucleus can be provided. Further, the use of this information is expected to contribute to the development of epoch-marking drugs.

/29

Sequence Table

(1) Name or designation of applicant: Helix Research Institute
(2) Title of Invention: METHODS FOR DETECTING AND ISOLATING NUCLEAR TRANSPORT PROTEINS
(3) Filing Number: H1-804DP1PCT
(4) Application Number:
(5) Filing Date:
(6) Name of Country and Numbers of Applications Relied on for Priority:

Japan Patent Application No. Hei 9-124795
Japan Patent Application No. Hei 9-309686

(7) Priority Date: April 28, 1998
October 24, 1998
(8) Number of Sequences: 26

Sequence number: 1
Length of sequence: 30
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Another nucleic acid Synthetic DNA
Sequence:

TTTGAATTCG CCAATTTTAA TCAAAGTGGG 30

Sequence number: 2
Length of sequence: 32
Form of sequence: Nucleic acid

/30

Number of strands: One
Topology: Straight chain
Type of Sequence: Another nucleic acid Synthetic DNA
Sequence:
TAGCATCTAT GACTTTTTGG GGCGTTCAAG TG 32

Sequence number: 3
Length of sequence: 342
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: cDNA to mRNA
Sequence characteristics:
Code denoting characteristics: domain
Position where present: 1..342
Method of determining characteristic: S
Sequence:

GCC AAT TTT AAT CAA AGT GGG AAT ATT GCT GAT AGC TCA TTG TCC TTC 48

Ala Asn Phe Asn Gln Ser Gly Asn Ile Ala Asp Ser Ser Leu Ser Phe

1 5 10 15

ACT TTC ACT AAC AGT AGC AAC GGT CCG AAC CTC ATA ACA ACT CAA ACA 96

Thr Phe Thr Asn Ser Ser Asn Gly Pro Asn Leu Ile Thr Thr Gln Thr

20 25 30

AAT TCT CAA GCG CTT TCA CAA CCA ATT GCC TCC TCT AAC GTT CAT GAT 144

Asn Ser Gln Ala Leu Ser Gln Pro Ile Ala Ser Ser Asn Val His Asp

35 40 45

AAC TTC ATG AAT AAT GAA ATC ACG GCT AGT AAA ATT GAT GAT GGT AAT 192

Asn Phe Met Asn Asn Glu Ile Thr Ala Ser Lys Ile Asp Asp Gly Asn

50 55 60

AAT TCA AAA CCA CTG TCA CCT GGT TGG ACG GAC CAA ACT GCG TAT AAC 240

Asn Ser Lys Pro Leu Ser Pro Gly Trp Thr Asp Gln Thr Ala Tyr Asn

65 70 75 80

GCG TTT GGA ATC ACT ACA GGG ATG TTT AAT ACC ACT ACA ATG GAT GAT 288

Ala Phe Gly Ile Thr Thr Gly Met Phe Asn Thr Thr Thr Met Asp Asp

85 90 95

GTA TAT AAC TAT CTA TTC GAT GAT GAA GAT ACC CCA CCA AAC CCA AAA 336

Val Tyr Asn Tyr Leu Phe Asp Asp Glu Asp Thr Pro Pro Asn Pro Lys

100 105 110

AAA GAG 342

Lys Glu

Sequence number: 4
Length of sequence: 609
Form of sequence: Nucleic acid
Number of strands: Two
Topology: Straight chain
Type of Sequence: cDNA to mRNA
Sequence characteristics:
Code denoting characteristics: CDS
Position where present: 1..606
Method of determining characteristic: S
Sequence:

ATG AAA GCG TTA ACG GCC AGG CAA CAA GAG GTG TTT GAT CTC ATC CGT 48

Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg

1 5 10 15

GAT CAC ATC AGC CAG ACA GGT ATG CCG CCG ACG CGT GCG GAA ATC GCG 96

Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala

20 25 30

CAG CGT TTG GGG TTC CGT TCC CCA AAC GCG GCT GAA GAA CAT CTG AAG 144

Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys

35 40 45

GCG CTG GCA CGC AAA GGC GTT ATT GAA ATT GTT TCC GGC GCA TCA CGC 192

Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg

50 55 60

GGG ATT CGT CTG TTG CAG GAA GAG GAA GAA GGG TTG CCG CTG GTA GGT 240

Gly Ile Arg Leu Leu Gln Glu Glu Glu Glu Gly Leu Pro Leu Val Gly

65 70 75 80

CGT GTG GCT GCC GGT GAA CCA CTT CTG GCG CAA CAG CAT ATT GAA GGT 288

Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly

85 90 95

CAT TAT CAG GTC GAT CCT TCC TTA TTC AAG CCG AAT GCT GAT TTC CTG 336

His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu

100 105 110

CTG CGC GTC AGC GGG ATG TCG ATG AAA GAT ATC GGC ATT ATG GAT GGT 384

Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly

115 120 125

GAC TTG CTG GCA GTG CAT AAA ACT CAG GAT GTA CGT AAC GGT CAG GTC 432

Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val

130 135 140

GTT GTC GCA CGT ATT GAT GAC GAA GTT ACC GTT AAG CGC CTG AAA AAA 480

Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys

145 150 155 160

CAG GGC AAT AAA GTC GAA CTG TTG CCA GAA AAT AGC GAG TTT AAA CCA 528

Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro

165 170 175

ATT GTC GTT GAC CTT CGT CAG CAG AGC TTC ACC ATT GAA GGG CTG GCG 576

Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala

180 185 190

GTT GGG GTT ATT CGC AAC GGC GAC TGG CTG TAA 609

Val Gly Val Ile Arg Asn Gly Asp Trp Leu

195 200

Sequence number: 5
Length of sequence: 10
Form of sequence: Amino acid
Topology: Straight chain
Type of Sequence: Peptide
Sequence
Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu
1 5 10

Sequence number: 6
Length of sequence: 30
Form of sequence: Nucleic acid

/34

Number of strands: One
Topology: Straight chain
Type of Sequence: synthetic DNA
Sequence
ACAGCTGCCA CCGATTGAGA GACTTACGTT 30

Sequence number: 7
Length of sequence: 30
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence

TGTCGACGGT GGCTAACTCT CTGAATGCAA 30

Sequence number: 8
Form of sequence: Nucleic acid
Number of strands: Two
Topology: Straight chain
Type of Sequence: cDNA to mRNA
Sequence characteristics:

Code denoting characteristics: CDS
Position where present: 1..1077
Method of determining characteristic: E

Sequence:

ATG AAA GCG TTA CAG CTG CCA CCG ATT GAG AGA CTT ACG TTA ACG GCC 48

Met Lys Ala Leu Gln Leu Pro Pro Ile Glu Arg Leu Thr Leu Thr Ala

1 5 10 15

AGG CAA CAA GAG GTG TTT GAT CTC ATC CGT GAT CAC ATC AGC CAG ACA 96

Arg Gln Gln Glu Val Phe Asp Leu Ile Arg Asp His Ile Ser Gln Thr

20 25 30

GGT ATG CCG CCG ACG CGT GCG GAA ATC GCG CAG CGT TTG GGG TTC CGT 144

Gly Met Pro Pro Thr Arg Ala Glu Ile Ala Gln Arg Leu Gly Phe Arg

35 40 45

TCC CCA AAC GCG GCT GAA GAA CAT CTG AAG GCG CTG GCA CGC AAA GGC 192

Ser Pro Asn Ala Ala Glu Glu His Leu Lys Ala Leu Ala Arg Lys Gly

50 55 60

GTT ATT GAA ATT GTT TCC GGC GCA TCA CGC GGG ATT CGT CTG TTG CAG 240

Val Ile Glu Ile Val Ser Gly Ala Ser Arg Gly Ile Arg Leu Leu Gln

65 70 75 80

GAA GAG GAA GAA GGG TTG CCG CTG GTA GGT CGT GTG GCT GCC GGT GAA 288

Glu Glu Glu Glu Gly Leu Pro Leu Val Gly Arg Val Ala Ala Gly Glu

85 90 95

CCA CTT CTG GCG CAA CAG CAT ATT GAA GGT CAT TAT CAG GTC GAT CCT 336

Pro Leu Leu Ala Gln Gln His Ile Glu Gly His Tyr Gln Val Asp Pro

100 105 110

TCC TTA TTC AAG CCG AAT GCT GAT TTC CTG CTG CGC GTC AGC GGG ATG 384

Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu Leu Arg Val Ser Gly Met

115 120 125

TCG ATG AAA GAT ATC GGC ATT ATG GAT GGT GAC TTG CTG GCA GTG CAT 432

Ser Met Lys Asp Ile Gly Ile Met Asp Gly Asp Leu Leu Ala Val His

130 135 140

AAA ACT CAG GAT GTA CGT AAC GGT CAG GTC GTT GTC GCA CGT ATT GAT 480
Lys Thr Gln Asp Val Arg Asn Gly Gln Val Val Val Ala Arg Ile Asp
145 150 155 160

GAC GAA GTT ACC GTT AAG CGC CTG AAA AAA CAG GGC AAT AAA GTC GAA 528
Asp Glu Val Thr Val Lys Arg Leu Lys Lys Gln Gly Asn Lys Val Glu
165 170 175

CTG TTG CCA GAA AAT AGC GAG TTT AAA CCA ATT GTC GTT GAC CTT CGT 576
Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro Ile Val Val Asp Leu Arg
180 185 190

CAG CAG AGC TTC ACC ATT GAA GGG CTG GCG GTT GGG GTT ATT CGC AAC 624
Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala Val Gly Val Ile Arg Asn
195 200 205

GGC GAC TGG CTG GAA TTC GCC AAT TTT AAT CAA AGT GGG AAT ATT GCT 672
Gly Asp Trp Leu Glu Phe Ala Asn Phe Asn Gln Ser Gly Asn Ile Ala
210 215 220

GAT AGC TCA TTG TCC TTC ACT TTC ACT AAC AGT AGC AAC GGT CCG AAC 720
Asp Ser Ser Leu Ser Phe Thr Phe Thr Asn Ser Ser Asn Gly Pro Asn
225 230 235 240

CTC ATA ACA ACT CAA ACA AAT TCT CAA GCG CTT TCA CAA CCA ATT GCC 768
Leu Ile Thr Thr Gln Thr Asn Ser Gln Ala Leu Ser Gln Pro Ile Ala
245 250 255

TCC TCT AAC GTT CAT GAT AAC TTC ATG AAT AAT GAA ATC ACG GCT AGT 816
Ser Ser Asn Val His Asp Asn Phe Met Asn Asn Glu Ile Thr Ala Ser
260 265 270

AAA ATT GAT GAT GGT AAT AAT TCA AAA CCA CTG TCA CCT GGT TGG ACG 864
Lys Ile Asp Asp Gly Asn Asn Ser Lys Pro Leu Ser Pro Gly Trp Thr

275 280 285

GAC CAA ACT GCG TAT AAC GCG TTT GGA ATC ACT ACA GGG ATG TTT AAT 912

Asp Gln Thr Ala Tyr Asn Ala Phe Gly Ile Thr Thr Gly Met Phe Asn

290 295 300

ACC ACT ACA ATG GAT GAT GTA TAT AAC TAT CTA TTC GAT GAT GAA GAT 960

Thr Thr Thr Met Asp Asp Val Tyr Asn Tyr Leu Phe Asp Asp Glu Asp

305 310 315 320

ACC CCA CCA AAC CCA AAA AAA GAG ATC TCT ATG GCT TAC CCA TAC GAT 1008

Thr Pro Pro Asn Pro Lys Lys Glu Ile Ser Met Ala Tyr Pro Tyr Asp

325 330 335

GTT CCA GAT TAC GCT AGC TTG GGT GGT CAT ATG GCC ATG GCG GCC GCT 1056

Val Pro Asp Tyr Ala Ser Leu Gly Gly His Met Ala Met Ala Ala Ala

340 345 350

CGA GTC GAC CTG CAG CCA AGC TAA 1080

Arg Val Asp Leu Gln Pro Ser

355

Sequence number: 9
Form of sequence: Nucleic acid
Number of strands: Two
Topology: Straight chain
Type of Sequence: cDNA to mRNA
Sequence characteristics:
Code denoting characteristics: CDS
Position where present: 1..1149

/38

Method of determining characteristic: S

Sequence:

ATG GCT AAG ATC TCT CCC GGG CTC GAG CTC ATG AAG AAG GGT ACT CAG 48
Met Ala Lys Ile Ser Pro Gly Leu Glu Leu Met Lys Lys Gly Thr Gln
1 5 10 15

CGT CTA TCC CGC CTG TTC GCC GCG ATG GCC ATT GCC GGG TTC GCC AGC 96
Arg Leu Ser Arg Leu Phe Ala Ala Met Ala Ile Ala Gly Phe Ala Ser
20 25 30

TAC TCC ATG GCC GCC GAC ACC ATC AAG ATC GCC CTG GCT GGC CCG GTC 144
Tyr Ser Met Ala Ala Asp Thr Ile Lys Ile Ala Leu Ala Gly Pro Val
35 40 45

ACC GGT CCG GTA GCC CAG TAC GGC GAC ATG CAG CGC GCC GGT GCG CTG 192
Thr Gly Pro Val Ala Gln Tyr Gly Asp Met Gln Arg Ala Gly Ala Leu
50 55 60

ATG GCA ATC GAA CAG ATC AAC AAG GCA GGC GGC GTG AAC GGC GCG CAA 240
Met Ala Ile Glu Gln Ile Asn Lys Ala Gly Gly Val Asn Gly Ala Gln
65 70 75 80

CTC GAA GGC GTG ATC TAC GAC GAC GCC TGC GAT CCC AAG CAG GCC GTG 288
Leu Glu Gly Val Ile Tyr Asp Asp Ala Cys Asp Pro Lys Gln Ala Val
85 90 95

GCG GTC GCC AAC AAG GTG GTC AAC GAC GGC GTC AAG TTC GTG GTC GGT 336
Ala Val Ala Asn Lys Val Val Asn Asp Gly Val Lys Phe Val Val Gly
100 105 110

CAT GTC TGC TCC AGC TCC ACC CAA CCC GCC ACC GAC ATC TAC GAA GAC 384
His Val Cys Ser Ser Ser Thr Gln Pro Ala Thr Asp Ile Tyr Glu Asp
115 120 125

GAA GGC GTG CTG ATG ATC ACC CCG TCG GCC ACC GCC CCG GAA ATC ACC 432

Glu Gly Val Leu Met Ile Thr Pro Ser Ala Thr Ala Pro Glu Ile Thr

130 135 140

TCG CGC GGC TAC AAG CTG ATC TTC CGC ACC ATC GGC CTG GAC AAC ATG 480

Ser Arg Gly Tyr Lys Leu Ile Phe Arg Thr Ile Gly Leu Asp Asn Met

145 150 155 160

CAG GGC CCG GTG GCC GGC AAG TTC ATC GCC GAA CGC TAC AAG GAC AAG 528

Gln Gly Pro Val Ala Gly Lys Phe Ile Ala Glu Arg Tyr Lys Asp Lys

165 170 175

ACC ATC GCG GTA CTG CAC GAC AAG CAG CAG TAC GGC GAA GGC ATC GCC 576

Thr Ile Ala Val Leu His Asp Lys Gln Gln Tyr Gly Glu Gly Ile Ala

180 185 190

ACC GAG GTG AAG AAG ACC GTG GAA GAC GCC GGC ATC AAG GTT GCC GTC 624

Thr Glu Val Lys Lys Thr Val Glu Asp Ala Gly Ile Lys Val Ala Val

195 200 205

TTC GAA GGC CTG AAC GCC GGC GAC AAG GAC TTC AAC GCG CTG ATC AGC 672

Phe Glu Gly Leu Asn Ala Gly Asp Lys Asp Phe Asn Ala Leu Ile Ser

210 215 220

AAG CTG AAG AAA GCC GGC GTG CAG TTC GTC TAC TTC GGC GGC TAC CAC 720

Lys Leu Lys Lys Ala Gly Val Gln Phe Val Tyr Phe Gly Gly Tyr His

225 230 235 240

CCA GAA ATG GGC CTG CTG CTG CGC CAG GCC AAG CAG GCC GGG CTG GAC 768

Pro Glu Met Gly Leu Leu Leu Arg Gln Ala Lys Gln Ala Gly Leu Asp

245 250 255

GCG CGC TTC ATG GGC CCG GAA GGG GTC GGC AAC AGC GAA ATC ACC GCG 816

Ala Arg Phe Met Gly Pro Glu Gly Val Gly Asn Ser Glu Ile Thr Ala

260 265 270

ATC GCC GGC GAC GCT TCG GAA GGC ATG CTG GCG ACC CTG CCG CGC GCC 864

Ile Ala Gly Asp Ala Ser Glu Gly Met Leu Ala Thr Leu Pro Arg Ala

275 280 285

TTC GAG CAG GAT CCG AAG AAC AAG GCC CTG ATC GAC GCC TTC AAG GCG 912

Phe Glu Gln Asp Pro Lys Asn Lys Ala Leu Ile Asp Ala Phe Lys Ala

290 295 300

AAG AAC CAG GAT CCG AGC GGC ATC TTC GTC CTG CCC GCC TAC TCC GCG 960

Lys Asn Gln Asp Pro Ser Gly Ile Phe Val Leu Pro Ala Tyr Ser Ala

305 310 315 320

GTC ACA GTG ATC GCC AAG GGC ATC GAG AAA GCC GGC GAG GCC GAT CCG 1008

Val Thr Val Ile Ala Lys Gly Ile Glu Lys Ala Gly Glu Ala Asp Pro

325 330 335

GAG AAG GTC GCC GAG GCC CTG CGC GCC AAC ACC TTC GAG ACT CCC ACC 1056

Glu Lys Val Ala Glu Ala Leu Arg Ala Asn Thr Phe Glu Thr Pro Thr

340 345 350

GGG AAC CTC GGG TTC GAC GAG AAG GGC GAC CTG AAG AAC TTC GAC TTC 1104

Gly Asn Leu Gly Phe Asp Glu Lys Gly Asp Leu Lys Asn Phe Asp Phe

355 360 365

ACC GTC TAC GAG TGG CAC AAG GAC GCC ACC CGG ACC GAG GTC AAG 1149

Thr Val Tyr Glu Trp His Lys Asp Ala Thr Arg Thr Glu Val Lys

370 375 380

TAA 1152

Sequence number: 10
Length of sequence: 12

/41

Form of sequence: Amino acid
Topology: Straight chain
Type of Sequence: Peptide
Sequence

Ser	Glu	Pro	Pro	Lys	Lys	Lys	Arg	Lys	Val	Glu	Thr
1				5					10		

Sequence number: 11
Length of sequence: 37
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
CTAGCGAGCC TCCAAAAAAG AAGAGAAAGG TCGAAAC 37

Sequence number: 12
Length of sequence: 37
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
GCTCGGAGGT TTTTTCTTCT CTTTCCAGCT TTGGTAC 37

Sequence number: 13

/42

Length of sequence: 419
Form of sequence: Nucleic acid
Number of strands: Two
Topology: Straight chain
Type of Sequence: Genomic DNA
Sequence

TCGACTGCTG TATATAAAAC CAGTGGTTAT ATGTACAGTA CTGCTGTATA TAAAACCAGT 60

GGTTATATGT ACAGTACGTC GAGGGAATCA AATTAACAAC CATAGGATGA TAATGCGATT 120

AGTTTTTTAG CCTTATTTCT GGGGTAATTA ATCAGCGAAG CGATGATTTT TGATCTATTA 180

ACAGATATAT AAATGCAAAA ACTGCATAAC CACTTTAACT AATACTTTCA ACATTTTCGG 240

TTTGTATTAC TTCTTATTCA AATGTAATAA AAGTATCAAC AAAAAATTGT TAATATACCT 300

CTATACTTTA ACGTCAAGGA GAAAAAACTA TAATGACTAA ATCTCATTCA GAAGAAGTGA 360

TTGTACCTGA GTTCAATTCT AGCGCAAAGG AATTACCAAG ACCATTGGCC GAAAAGTGC 419

Sequence number: 14
Length of sequence: 12
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
AATTGACCAC CC 12

Sequence number: 15
Length of sequence: 12
Form of sequence: Nucleic acid

/43

Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
CTGGTGGGTT AA 12

Sequence number: 16
Length of sequence: 25
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
CTAGCTTGGG TGGAATTCAT ATGGC 25

Sequence number: 17
Length of sequence: 24
Form of sequence: Nucleic acid

Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
GAACCCACCT TAAGTATACG GTAC 24

Sequence number: 18
Length of sequence: 11

/44

Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
CTGCATGCAC C 11

Sequence number: 19
Length of sequence: 14
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
ATGGACGTAC GTGG 14

Sequence number: 20
Length of sequence: 32
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
CTATTCGATG ATGAAGATAC CCCACCAAAC CC 32

Sequence number: 21

/45

Length of sequence: 30
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
GAAATTCGCC CGGAATTAGC TTGGCTGCAG 30

Sequence number: 22

Length of sequence: 32
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
CTATTCGATG ATGAAGATAC CCCACCAAAC CC 32

Sequence number: 23
Length of sequence: 30
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
GAAATTCGCC CGGAATTAGC TTGGCTGCAG 30

/46

Sequence number: 24
Length of sequence: 32
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
TTTCAATTGG AATGGACGAA CTGTTCCCCC TC 32

Sequence number: 25
Length of sequence: 35
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGG 35

Sequence number: 26
Length of sequence: 35
Form of sequence: Nucleic acid
Number of strands: One
Topology: Straight chain
Type of Sequence: Other nucleic acid synthetic DNA
Sequence
TTTGAATTCT AATGATGTTC TCGGGTTTCA ACGCG 35

/47

Claims

(1) A method of detecting the nuclear transportability of a peptide coded for by test DNA, characterized in that fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is introduced into a eukaryotic host having in the nucleus thereof a promoter region that is activated by the binding of transcription factor and a reporter gene connected downstream from said promoter region, and in that expression of the reporter gene is detected.

(2) The method of claim (1) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA bonding domain, and a transcription activation domain.

(3) The method of claim (1) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and a GAL4 transcription activation domain, and wherein the promoter region that is activated by binding of the transcription factor is the promoter region of a GAL1 gene in which the operator sequence has been replaced with the LexA operator sequence.

(4) The method of claim (3) wherein said nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5.

(5) The method of any of claims (1)-(4) wherein said reporter gene is the LEU2 and/or β-galactosidase gene.

(6) A method of isolating DNA coding for a peptide having nuclear transportability characterized in that fused DNA of DNA coding for transcription factor not having nuclear transportability and test DNA is introduced into a eukaryotic host having in its nucleus a promoter region that is activated by binding of said transcription factor and a reporter gene connected downstream from said promoter region, in that expression of the reporter gene is detected, and in that test DNA is isolated from a eukaryotic host in which said expression has been detected.

(7) The method of claim (6) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA binding domain, and a transcription activation domain.

/48

(8) The method of claim (6) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and a GAL4 transcription activation domain, and wherein said promoter region that is activated by binding of the transcription factor is the

promoter region of a GAL1 gene in which the operator sequence has been replaced with the LexA operator sequence.

(9) The method of claim (8) wherein said nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5.

(10) The method of any of claims (6)-(9) wherein said reporter gene is the LEU2 and/or β-galactosidase gene.

(11) A vector having an incorporation site of test DNA adjacent to DNA coding for transcription factor not having nuclear transportability.

(12) The vector of claim (11) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA binding domain, and a transcription activation domain.

(13) The vector of claim (11) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and the GAL4 transcription activation domain.

(14) The vector of claim (13) wherein said nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5.

(15) A kit comprising:

(1) a vector having an incorporation site for test DNA adjacent to DNA coding for transcription factor not having nuclear transportability; and

(2) a eukaryotic host having in its nucleus an expression unit comprising a promoter region activated by binding of said transcription factor and a reporter gene connected downstream from said promoter region.

(16) The kit of claim (15) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, a DNA binding domain, and a transcription activation domain.

/49

(17) The kit of claim (15) wherein said transcription factor not having nuclear transportability is a fused protein comprising a nuclear export signal, LexA protein, and a GAL4 transcription activation domain; wherein said promoter region that is activated by binding of said transcription factor is the promoter region of a GAL1 gene in which the operator sequence has been replaced with the LexA operator sequence; and wherein said eukaryotic host is yeast.

(18) The kit of claim (17) wherein said nuclear export signal is a peptide comprising the amino acid sequence described in sequence number 5.

(19) The kit of any of claims (15)-(18) wherein said reporter gene is the LEU2 and/or β-galactosidase gene.

1/8

Keys to Fig. 1:
(1) Ampicillin

2/8

Keys to Fig. 2:
(1) Ampicillin

3/8

Keys to Fig. 3:
(1) Ampicillin

4/8

Keys to Fig. 4:
(1) Ampicillin

5/8

Keys to Fig. 7:
(1) Ampicillin

図1

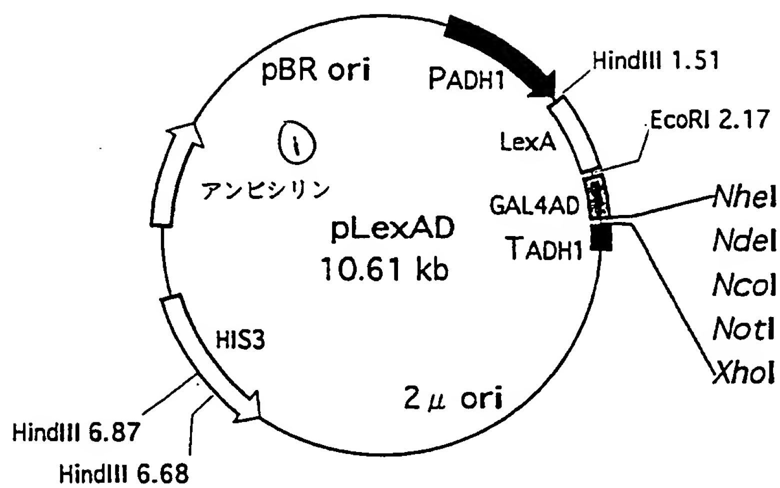

pBR ori
PADH1
HindIII 1.51
EcoRI 2.17
LexA
アンピシリン
GAL4AD
TADH1
NheI
NdeI
NcoI
NotI
XhoI
pLexAD
10.61 kb
HIS3
2μ ori
HindIII 6.87
HindIII 6.68

図2

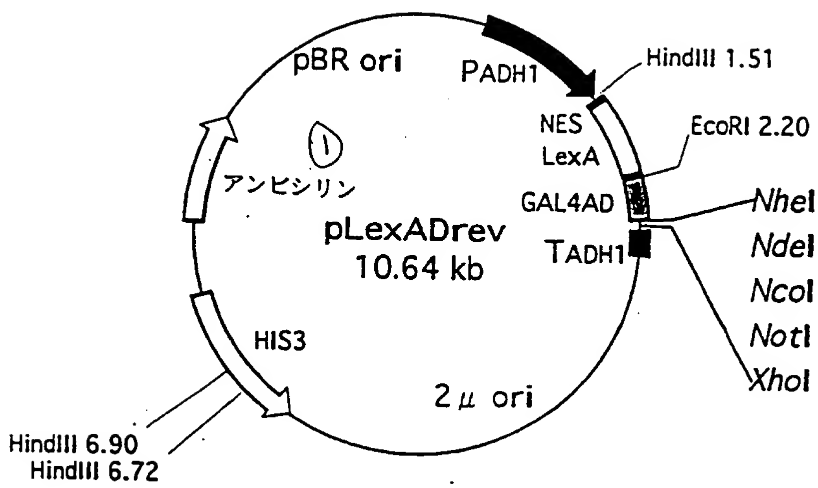

pBR ori
PADH1
HindIII 1.51
NES
LexA
EcoRI 2.20
①
アンピシリン
GAL4AD
NheI
pLexADrev
10.64 kb
TADH1
NdeI
NcoI
NotI
HIS3
XhoI
2μ ori
HindIII 6.90
HindIII 6.72

図3

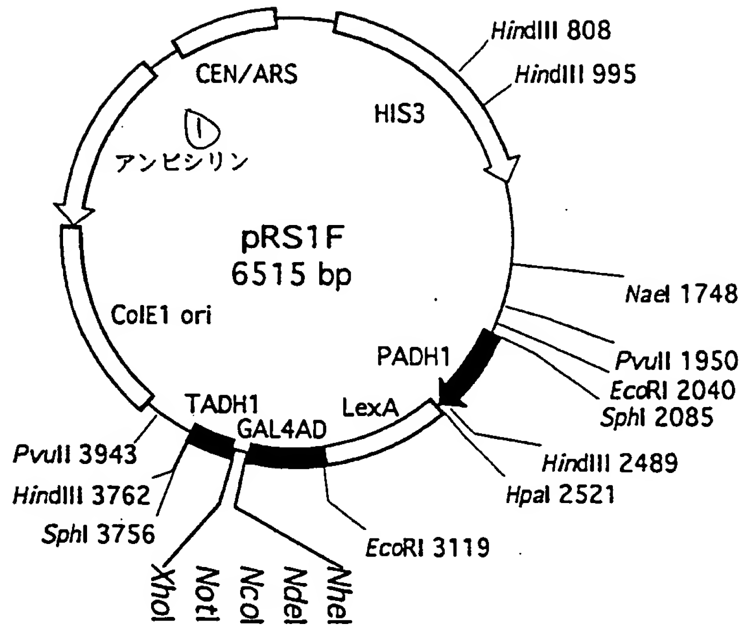

HindIII 808
HindIII 995
CEN/ARS
HIS3
1
アンピシリン
pRS1F
6515 bp
NaeI 1748
ColE1 ori
PADH1
PvuII 1950
EcoRI 2040
SphI 2085
TADH1
LexA
GAL4AD
PvuII 3943
HindIII 2489
HindIII 3762
HpaI 2521
SphI 3756
EcoRI 3119
XhoI
NotI
NcoI
NdeI
NheI

図4

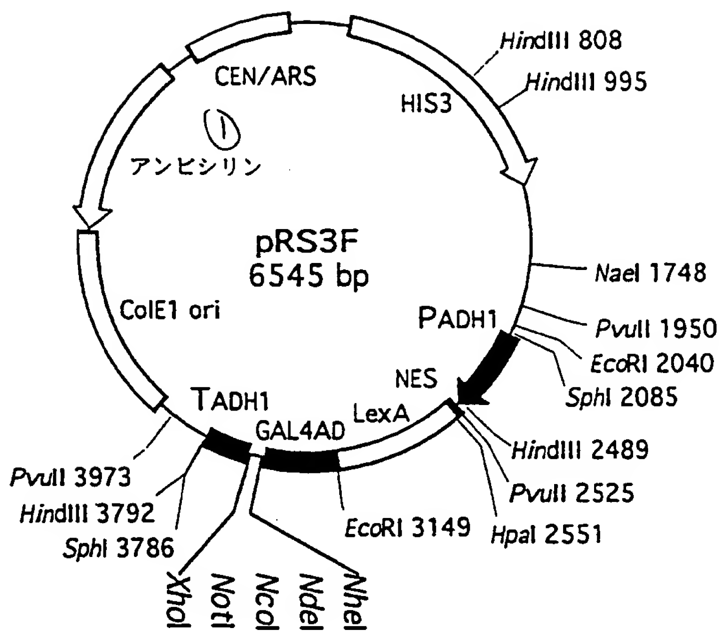

CEN/ARS
HIS3
HindIII 808
HindIII 995
①
アンピシリン
pRS3F
6545 bp
NaeI 1748
ColE1 ori
PADH1
PvuII 1950
EcoRI 2040
NES
SphI 2085
TADH1
LexA
GAL4AD
HindIII 2489
PvuII 2525
HpaI 2551
PvuII 3973
HindIII 3792
SphI 3786
EcoRI 3149
XhoI
NotI
NcoI
NdeI
NheI

図5

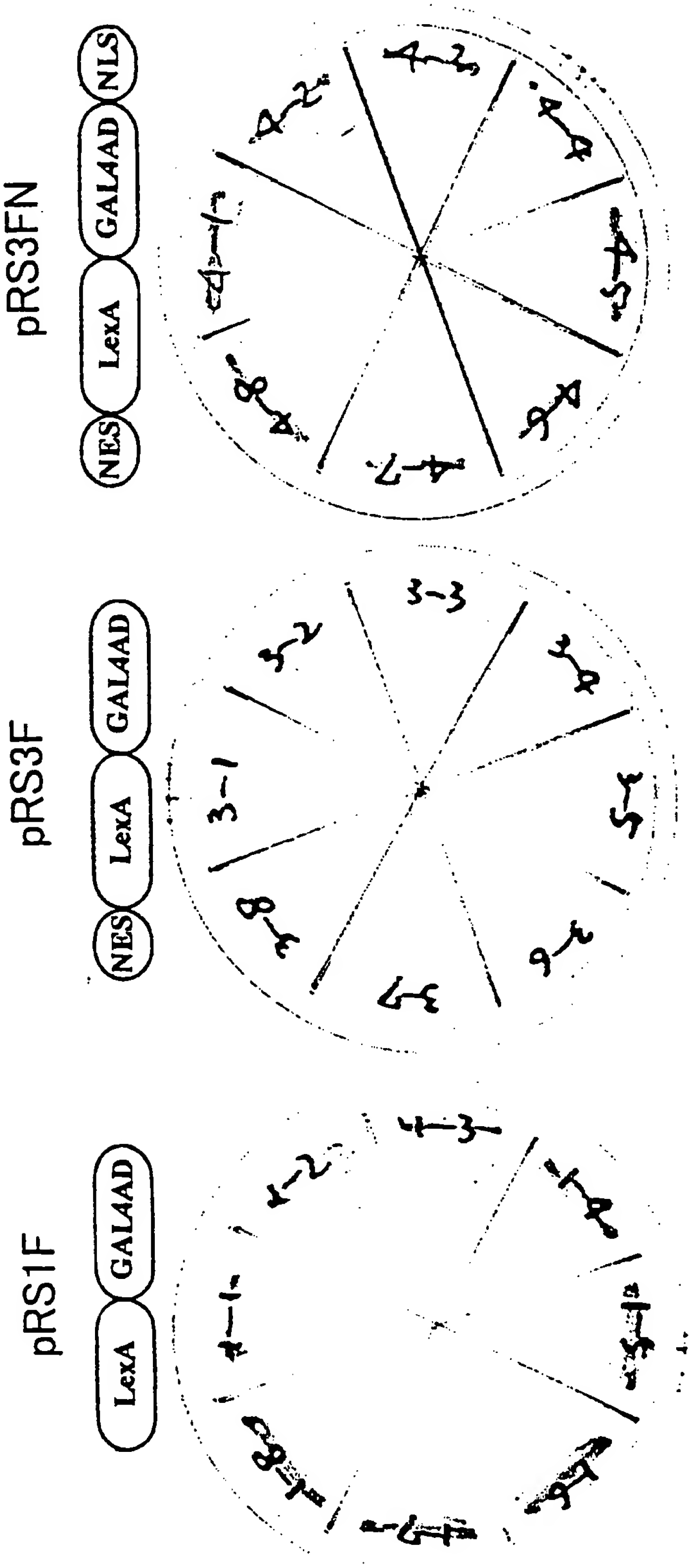

pRS3FN
NES
LexA
GALAAD
NLS
pRS3F
NES
LexA
GALAAD
3-1
3-3
pRS1F
LexA
GALAAD

図6

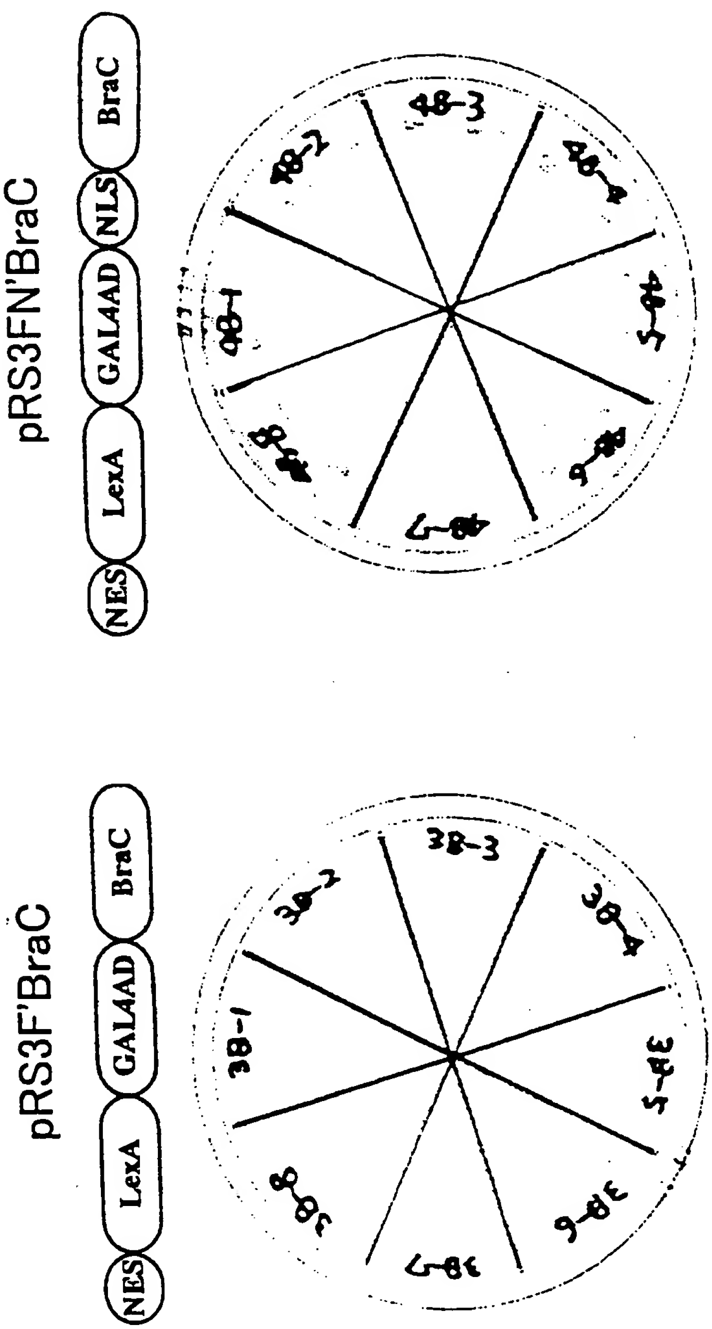

pRS3FN'BraC
NES
LexA
GALAAD
NLS
BraC
4B-1
4B-2
4B-3
4B-4
4B-5
4B-6
4B-7
4B-8
pRS3F'BraC
NES
LexA
GALAAD
BraC
3B-1
3B-2
3B-3
3B-4
3B-5
3B-6
3B-7
3B-8

図7

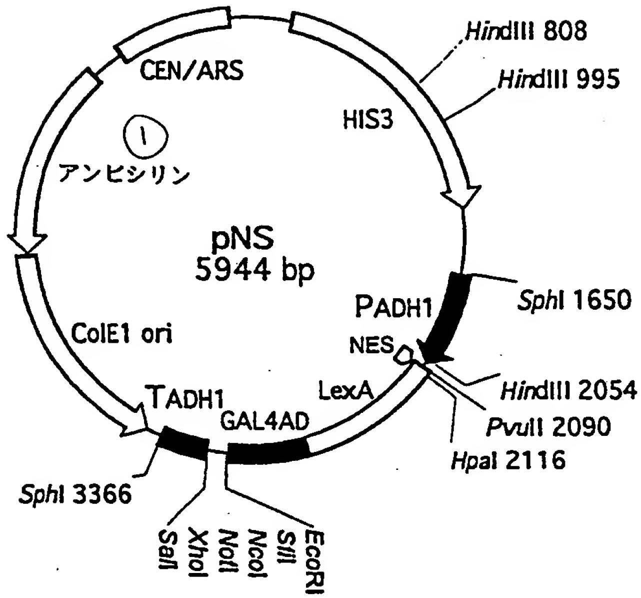
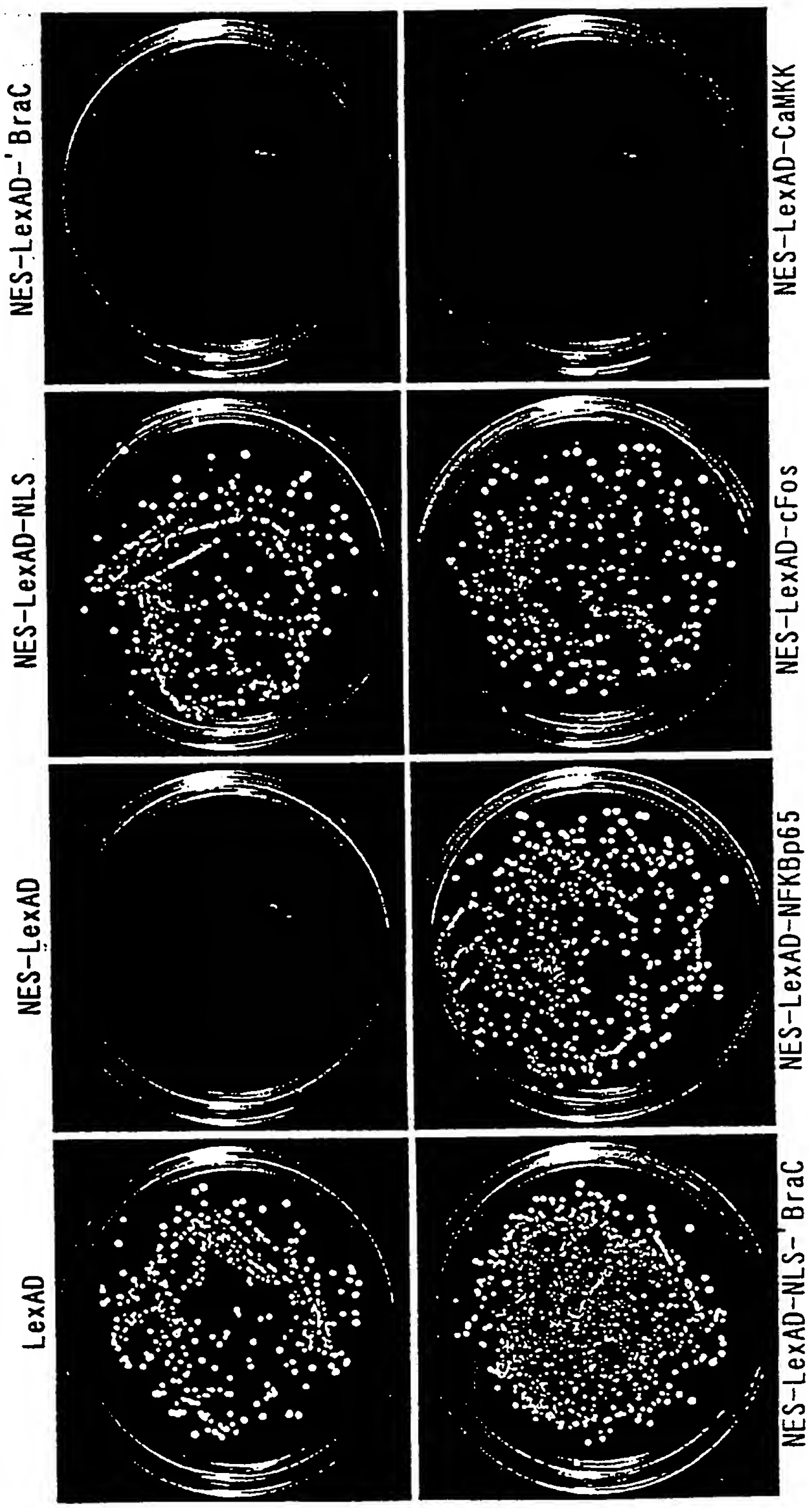

HindIII 808
HindIII 995
CEN/ARS
HIS3
1
アンピシリン
pNS
5944 bp
PADH1
SphI 1650
ColE1 ori
NES
HindIII 2054
PvuII 2090
HpaI 2116
TADH1
LexA
GAL4AD
SphI 3366
SalI
XhoI
NotI
NcoI
SfiI
EcoRI

図 8


NES-LexAD-'BraC
NES-LexAD-CaMKK
NES-LexAD-NLS
NES-LexAD-cFos
NES-LexAD
NES-LexAD-NFKBp65
LexAD
NES-LexAD-NLS-'BraC